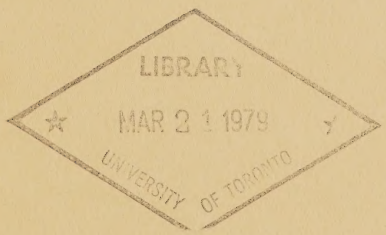
Ontario

Commission on Freedom of Information and Individual Privacy

Research and Statistical Uses of Ontario Government Personal Data



Research Publication 5

RESEARCH AND STATISTICAL USES OF ONTARIO GOVERNMENT PERSONAL DATA

by David H. Flaherty
Professor of History
University of Western Ontario

Research Publication 5

Prepared for the
Commission on Freedom of Information
and Individual Privacy

January, 1979

ISBN 0-7743-3229-8

Ontario

D. Carlton Williams, Ph.D.
Chairman
Dorothy J. Burgoyne, B.A.
G. H. U. Bayly, M.Sc.F.
Members
W. R. Poole, Q.C.
Counsel
J. D. McCamus, L.L.M.
Director of Research
Hon. J. C. McRuer, O.C.
Consultant
Doris E. Wagg
Registrar

Commission
on
Freedom of Information
and
Individual Privacy

416/598-0411

180 Dundas Street West
22nd Floor
Toronto Ontario
M5G 1Z8

FOREWORD

The Commission on Freedom of Information and Individual Privacy was established by the government of Ontario in March, 1977, to "study and report to the Attorney General of Ontario on ways and means to improve the public information policies and relevant legislation and procedures of the government of Ontario, and to examine:

1. Public information practices of other jurisdictions in order to consider possible changes which are compatible with the parliamentary traditions of the government of Ontario and complementary to the mechanisms that presently exist for the protection of the rights of individuals;

2. The individual's right of access and appeal in relation to the use of government information;

3. The categories of government information which should be treated as confidential in order to protect the public interest;

4. The effectiveness of present procedures for the dissemination of government information to the public;

5. The protection of individual privacy and the right of recourse in regard to the use of government records."

To the best of our knowledge it is the only Commission of its kind whose mandate ebraces both freedom of information and individual privacy. The views of the public were embodied in the briefs submitted and in the series of hearings held in ten communities, and covering both Northern and Southern Ontario. In response to public demand, three sets of hearings, widely separated in time, were held in Toronto.

The views of the scholars and experts in the field are to be found in the present series of research reports of which this is number 5. These, together with the briefs submitted, constitute the backbone of our findings: the stuff out of which our Report will be made. Many of these stand in their own right as documents of importance to this field of study; hence our decision to publish them immediately.

It is our confident expectation that they will be received by the interested public with the same interest and enthusiasm they generated in us. Many tackle problem areas never before explored in the context of freedom of information and individual privacy in Canada. Many turn up facts, acts, policies and procedures hitherto unknown to the general public.

In short, we feel that the Commission has done itself and the province a good turn in having these matters looked into and that we therefore have an obligation in the name of freedom of information to make them available to all who care to read them.

It goes without saying that the views expressed are those of the authors concerned; none of whom speak for the Commission.

D. C. Williams
Chairman

PREFACE

The inherent tension between the two branches of this Commission's mandate -- freedom of information and the protection of privacy -- is brought into sharp relief when attention is drawn to the vast quantity of personal information about individual citizens which is collected by the government of Ontario. A critic of the Ontario medical health insurance system, for example, may take a very different view of the question of public access to records generated by that system than would a hospital patient whose medical history is therein revealed, or a medical practitioner whose personal income is, for the most part, reflected in receipts from the insurance fund. In such cases, as one wag has remarked, "One man's freedom of information is another man's invasion of privacy."

It is evident that the two very different policy objectives of public access to government information and privacy protection ought not be considered in isolation. There is a danger that forceful advocates of citizen access to government documents will lose sight of the risks to individual privacy posed by unduly aggressive freedom of information laws. On the other hand, an overly zealous pursuit of the privacy protection objective might result in a reduction of access to government-held information to an extent which would be contrary to the public interest.

This tension between the access and privacy protection objectives is nowhere more dramatically illustrated than in the area of information practice discussed in this paper, the use of sensitive personal information for research and statistical purposes. Few would contest that there is a strong public interest to be served in granting the research community access to government-held information about individuals which is relevant to legitimate research objectives. Those who remain skeptical on this point may find that Professor Flaherty's account of recent attempts to conduct medical and epidemiological research offers compelling evidence of the validity of this proposition.

It is undeniable, however, that the conduct of medical and other research drawing on sensitive personal information can pose a grave risk to privacy interests unless precautions are taken to safeguard the confidentiality of data of this kind. This much is obvious. It is rather more difficult to determine the precise nature of the precautions which should be followed and the extent to which uniformity in such matters is practicable or ought to be encouraged through the imposition of legislated standards.

As Professor Flaherty's paper indicates, a closer examination of the problems associated with research uses of personal information yields a number of equally challenging questions. If the research community

is currently encountering difficulty in gaining access to government-held personal information, is legislative support of some kind to ensure or foster access appropriate? Should government agencies ever be allowed to use personal information, initially gathered for a research or statistical use, to accomplish an administrative objective of some kind? Conversely, under what circumstances should information gathered for an administrative purpose be used by government or by outside researchers for legitimate research purposes? Under what terms and conditions ought linkage of information from a variety of different government-held personal files or data banks be permitted to accomplish research objectives? What are the implications of research needs for the current national debate on the future use of the Social Insurance Number?

Professor Flaherty's paper provides the Commission and interested members of the public with a factual base drawn from the Ontario experience against which questions of this kind may be considered. More than this, however, Professor Flaherty has offered suggestions for the resolution of some of these issues and, indeed, has made specific recommendations with respect to the information practices of a number of Ontario research and statistical agencies.

Professor Flaherty is ideally suited to undertake a study of this topic for the Commission. As a professional historian, Professor Flaherty has long had an interest in the concept of personal privacy (see his Privacy in Colonial New England, 1972). More recently, his research activities have been directly focused on privacy problems associated with research and statistical uses of personal data. As he indicates in his Introduction, Professor Flaherty has recently completed a lengthy inquiry into the data collection and dissemination activities of national statistical agencies in five western countries. Professor Flaherty is a member of the Department of History of the University of Western Ontario.

It should be noted that the interviews and other inquiries which form the basis of the descriptive portions of this paper were conducted by Professor Flaherty in June and July of 1978. It is entirely possible, then, that the policies and practices of the agencies reviewed by Professor Flaherty may have altered in the period prior to publication of this report.

The Commission has resolved to make available to the public its background research papers in the hope that they might stimulate public discussion. Those who wish to communicate their views in writing to the Commission are invited to write to us at the following address:

Registrar
Commission on Freedom of Information
and Individual Privacy
180 Dundas Street West, 22nd Floor
Toronto, Ontario M5G 1Z8

It should be emphasized, however, that the views expressed in this paper are those of the author and that they deal with questions on which the Commission has not yet reached a final conclusion.

Particulars of other research papers which have been published to date by the Commission are to be found on page 188.

John D. McCamus
Director of Research

RESEARCH AND STATISTICAL USES OF
ONTARIO GOVERNMENT PERSONAL DATA

TABLE OF CONTENTS

RESEARCH AND STATISTICAL USES OF
ONTARIO GOVERNMENT PERSONAL DATA

LIST OF TABLES

LIST OF TABLES (cont'd)

RESEARCH AND STATISTICAL USES OF
ONTARIO GOVERNMENT PERSONAL DATA

TABLE OF ABBREVIATIONS

ARF	Alcholism and Drug Addiction Research Foundation
CSS	Central Statistical Services
DDEB	Data Development and Evaluations Branch Information Systems Division Ministry of Health
HMRI	Hospital Medical Records Institute
MGS	Ministry of Government Services
MOH	Ministry of Health
OCTRF	Ontario Cancer Treatment and Research Foundation
OEC	Ontario Economic Council
OHIP	Ontario Health Insurance Plan
SMAC	Systems Management and Co-ordination Branch Information Systems Division Ministry of Health
TEIGA	Ministry of Treasury, Economics and Intergovernmental Affairs
UPI	Unique Personal Identifier

INTRODUCTION

Various levels of government in advanced industrial societies collect large amounts of data on individuals (microdata) in the course of carrying out their responsibilities. At the provincial level in Canada, most of the data originates with the process of administering a government program, such as hospitalization insurance. Such administrative data have varying levels of sensitivity. Data are also collected for purely statistical purposes.

The basic thrust of contemporary movements for data protection insists on a variety of very important and relevant themes, which are equally applicable to the data collection activities of the government of Ontario. These themes were well summarized in the 1977 report of the American Privacy Protection Study Commission, which listed the following objectives for national policy-making in an information society: minimizing intrusiveness in data collection; maximizing fairness in recording and storing personal data; and legitimizing expectations of confidentiality for data.[1]

1 Privacy Protection Study Commission, Personal Privacy In An Information Society (Washington, D.C., 1977).

But there is another facet to the issue of data collection by government agencies. Most information gathering occurs for a particular purpose that is thought to be in the public interest. For example, administrative data may determine eligibility for various types of personal benefits. Data protection thus should not unduly hinder the conduct of legitimate government activities. Another risk is that advocates of privacy and data protection will take an absolutist position on the use of data for purposes other than those for which they were collected. The privacy lobby, for example, could conceivably cut off the use of data for research and statistical purposes because of the risks to personal privacy allegedly involved. Research and statistical or scientific uses of data are defined as uses which do not directly affect a particular person on the basis of the specific data in question. These are in sharp contrast to administrative or regulatory uses of data, which directly affect a person in one way or another.

The issue of protecting versus disseminating data is at the heart of the dilemma over privacy and freedom of information. The formulation of essential policies for the protection of personal data could have the unfortunate result of restricting the flow of information for legitimate and important societal purposes. Thus the goals of protection and dissemination have to be balanced during the formulation of measures to protect

the confidentiality of data. Balancing the appropriate mechanisms has been the major theme pursued in the preparation of this Working Paper.

This Working Paper has been substantially influenced by the methodology and findings of the forthcoming book by David H. Flaherty on Privacy and Government Data Banks. An International Perspective.[2] The companion volume compiled by D.H. Flaherty, E.H. Hanis, and S. Paula Mitchell will be entitled Privacy and Access to Government Data for Research. An International Bibliography.[3] Both volumes are products of a research project funded by the Ford Foundation from 1974 to 1978. The monograph surveys data collection and data dissemination activities by national government agencies in five countries.

Although Privacy and Government Data Banks focuses on scientific uses of data, the agencies surveyed in the volume include both administrative and statistical data bases. In the United Kingdom, for example, the government agencies studied include the Office of Population Censuses and Surveys, the Central Statistical Office, the Department of Health and Social Security,

2 Science Associates/International, Inc., New York, forthcoming, 1978.

3 Ibid.

the Board of Inland Revenue, and the Home Office. This English material has furnished instructive background for surveying agencies and issues in Ontario.

For each of the Ontario agencies examined in this Working Paper, the following types of questions were pertinent: what personal data does an agency of government collect from respondents or from administrative sources? What general policies and legal protections exist to regulate data protection and dissemination? How are these policies formulated? What mechanisms for data dissemination exist? These concerns are balanced by a discussion of various types of existing and future needs for access to data for research and statistical purposes, drawing upon particular examples of current research.

The basic goal of this Working Paper has been to generate a set of principles and practices for data protection and dissemination that will be generally applicable to a wide variety of Ontario agencies involved with personal data. The recommendations are generally patterned on the Bellagio Principles on the Use of Government Microdata for Research and Statistical Purposes, which were formulated at the Bellagio Conference in August, 1977, and the conclusions and recommendations of Privacy and Government Data Banks.[4]

4 The Bellagio Conference is discussed in Appendix I (p. 182). The conclusions of Privacy and Government Data Banks are summarized in Table IX.

Several Ontario government agencies were selected for intensive review in order to examine the questions raised above. Two of the chosen agencies are primarily statistical in character: the Office of the Registrar General in the Ministry of Consumer and Commercial Relations, and Central Statistical Services in the Ministry of Treasury, Economics, and Intergovernmental Affairs. An initial focus on these particular branches of government permitted the application of knowledge about Statistics Canada and the national statistical system acquired during research for Privacy and Government Data Banks.

A second major area of inquiry has been the various components of the Ontario Ministry of Health and related health research foundations that collect and disseminate health-related data on individuals. This study focuses on sources of personal data in this Ministry with potential or actual research and statistical applications. Not only are the uses of OHIP data a sensitive topic at present from the point of view of confidentiality, but health data have some of the best potential applications for purposes of scientific research. Hence this inquiry also interviewed a number of users of such data within and outside the Ontario government, whose experience as users is described. Striking a reasonable balance between the protection and use of personal data is a major goal of this Working Paper. In addition, the promising field of record linkage for scientific purposes is

a particular field of emphasis in these pages, since it is currently one of the most sensitive issues confronting custodians of government data in English-speaking countries.

Several caveats are in order for readers of the following material. The Working Paper focuses heavily on legitimate uses of Ontario government personal data for research and statistical purposes. The focus is not on even more fundamental issues such as whether data should have been collected in the first instance, and whether the basic confidentiality of the data is adequately protected during administrative uses. The latter is a direct concern for health-related data of the Royal Commission of Inquiry into the Confidentiality of Health Records in Ontario headed by Justice Horace Krever. This study reviews existing government sources of personal microdata and describes how these are or should be used for scientific purposes under controlled conditions. Yet it must be emphasized that it is also a basic assumption of this Working Paper that personal data should be collected and used in a manner that provides for basic protection of the right of individual privacy of citizens. All uses of data discussed in this Working Paper are research and statistical in character, wherein there should be no direct action for or against a particular person, except in the unusual circumstances where treatment or action are part of the research program or protocol and approved as such by the appropriate

custodians of data. The thesis of the following pages is that existing personal data should be used as much as possible to promote scientific understanding of important contemporary problems.

CHAPTER I

OFFICE OF THE REGISTRAR GENERAL

The Office of the Registrar General is located in the Ministry of Consumer and Commercial Relations. Although the Minister is the Registrar General, the direct supervision of the Office is delegated to the Deputy Registrar General under the Vital Statistics Act.[1] The Office of the Registrar General is responsible for the collection and custody of all personal records required under the Vital Statistics Act, which dates back to 1948. It thus has both administrative and statistical responsibilities. It issues certificates of births, marriages, and deaths, records vital events, and furnishes statistical data in published and unpublished form.

The Office of the Registrar General is legally required to maintain a large amount of personal data on Ontario residents. There are seven separate annual series recording identifiable individual data on births, deaths, marriages, still-births, adoptions, divorces, and changes of name. Each entry in each

1 R.S.O., 1970, c. 483, as amended by S.O., 1971, c. 98, Sched., par. 35, s. 4(1).

series acquires a specific consecutive number. Each series is indexed on an annual basis.[2] The annual registrations of Ontario births, deaths, and marriages total approximately 126,000, 62,000, and 73,000 items respectively.[3] Data are stored in volumes, on punch cards, and/or on computer tapes. There are proposals for increasing automation of the Office of the Registrar General.

The data collected on individual episodes of mortality furnish some sense of the extensive personal information collected and stored by the Registrar General. Information on mortality is collected by means of three separate forms. Table I contains a comparative analysis of the personal data contained in the various documents. The "statement of death" prepared by a funeral director reports approximately twenty items of information about the deceased.[4] The statement of death contains a great deal of valuable information from the point of view of the health researcher. The same is also true for the

2 R.S.O., 1970, c. 483, s. 2, 3; Regulation 820, s. 63, 64.

3 Ministry of Treasury, Economics and Intergovernmental Affairs, Catalogue of Statistical Files in the Ontario Government 1977 (Toronto, 1977), pp. CR.12 - CR.13.

4 Vital Statistics Act, R.S.O., 1970, c. 483, s. 17(4); Regulation 820, Form 15.

TABLE I: CONTENTS OF MORTALITY RECORDS
OFFICE OF THE REGISTRAR GENERAL

Item	"Statement of Death" (Form 15)	"Death Certificate" (Form 29)	"Medical Certificate of Death" (Form 16)
Name of Deceased	x	x	x
Registration Number	x	x	x
Place of Death	x	x	x
Date of Death	x	x	x
Length of Residence in Place of Death	x		
Length of Residence in Ontario	x		
Length of Residence in Canada	x		
Permanent Residence	x		
Cause of Death			x
Sex	x	x	x
Citizenship	x		
Birthplace	x		
Birthdate	x		
Age	x		
Trade, Profession	x		
Years in Occupation	x		
Type of Business	x		
Marital Status	x	x	
Name of Father	x		
Name of Mother	x		
Birthplace of Father	x		
Birthplace of Mother	x		
Name of Husband/Maiden Name of Wife	x		

Source: Regulations Under the Vital Statistics Act. Revised Statutes of Ontario, 1970. Regulation 820. December, 1971 (Toronto, 1971).

"medical certificate of death," which is supplied by a legally qualified medical practitioner.[5] Form 16 furnishes detailed information concerning the cause of death; forms 15 and 16 are in fact filed together as one record in the Office. Finally, there is a "death certificate," which contains a limited range of identifiable information about the deceased.[6]

Confidentiality and Data Dissemination

The Vital Statistics Act grants broad discretionary power to the Registrar General to "collate, publish and distribute such statistical information regarding the births, marriages, deaths, still-births, adoptions, divorces, and changes of names registered during any period as he may consider to be necessary and in the public interest."[7] The meaning of "statistical information" is in part established in Section 48(2) of the Act, which is cited below. No criteria are established for the determination of "necessity" and the "public interest." Although the inclusion of definitions or explications of such terms could have been restrictive, users now run the risk of changing attitudes on the part of custodians.

5 R.S.O., 1970, c. 483, s. 17(3); Regulation 820, Form 16.

6 Regulation 820, Form 29.

7 R.S.O., 1970, c. 483, s. 3(4).

The basic provision on secrecy in the Act is an indirect promise of confidentiality to respondents:

> 48(1)
> "No division registrar, sub-registrar, funeral director or person employed in the service of Her Majesty shall communicate or allow to be communicated to any person not entitled thereto any information obtained under this Act, or allow any such person to inspect or have access to any records containing information obtained under this Act.
> (2)
> Nothing in subsection 1 prohibits the furnishing and publication of information of a general statistical nature that does not disclose information about any individual person."

The Deputy Registrar General interprets Section 48(1) as applying to all persons employed by the Office of the Registrar General. There is no specific definition of who is "entitled" to information under Section 48(1). The regulations under the Act do specify rights of access for various types of government officials.[8] Over the course of many years, the Deputy Registrar General has generally determined that the individuals "entitled" to direct access to information are the persons whose records are being required or their immediate families. This is in accord with the 1972 revision to the Coroners Act, which directs that coroners should make information recorded by them about specific cases "available to the spouse, parents, children, brothers and sisters of the deceased and to his personal representative, upon request."[9] It is evident that researchers do not fit this

8 Regulation 820, s. 66.

9 S.O., 1972, c. 98, s. 16(2).

particular set of categories; the legitimacy of access by qualified researchers should be recognized under existing regulations. The Vital Statistics Act specifies that "any person contravening any of the provisions of Section 48 is guilty of an offense and on summary conviction is liable to a fine of not more than $200."[10] There has never been a known episode requiring application of this section.

The Act has a variety of provisions concerning access to data. Only those pertaining to mortality data will be reviewed here, since this is the type of information most commonly required by health researchers. In general, any person who applies, pays the prescribed fee, and "satisfies the Registrar General as to his reason for requiring it," may have a search made for the registration of any death in the indexes kept in the Office of the Registrar General. A request cannot be for frivolous purposes. The only information that can be disseminated under this power of search shall concern the existence or otherwise of the registration and the registration number if registered.[11] A more relevant provision concerning access for researchers would be the ability to acquire copies of death certificates: "Upon application and upon payment of the prescribed fee, any person

10 R.S.O., 1970, c. 483, s. 52.

11 R.S.O., 1970, c. 483, s. 43(1)(3).

may obtain from the Registrar General a death certificate in respect of any death of which there is a registration in his Office."[12] In fact, a "death certificate" contains only limited information about the deceased, as indicated in Table I.[13] Copying of the "medical certificate of death" is controlled by the Vital Statistics Amendment Act of 1973: "No person shall make a copy or a duplicate of the medical certificate of death, nor shall any person receive a copy of the certificate, except as authorized by this or any other Act, or the regulations made thereunder."[14]

The Vital Statistics Act does authorize the Lieutenant Governor in Council to make regulations "designating the persons who may have access to or may be given information from the records in the Registrar General's Office or in a Division Registrar's Office, and prescribing an oath of secrecy to be taken by such persons."[15] None of the existing regulations directly pertain to access for research and statistical purposes. The regulations do prescribe the forms for two separate oaths of secrecy, neither of which make any specific provision of penalties for violation of the oath.

12 R.S.O., 1970, c. 483, s. 39(2).

13 R.S.O., 1970, c. 483, s. 38(2).

14 S.O., 1973, c. 114, s. 5(5).

15 R.S.O., 1970, c. 483, s. 54(J).

The first can be reviewed in Table II. Under the oath of secrecy prescribed in form 31 the individual swears as follows:

> "I will hold secret and will not disclose to any person any information given me from the records in the Registrar General's Office or in any Division Registrar's Office or obtained from those records by reason of my access thereto, except to the Director of Industrial Hygiene or the Medical Statistician of the Department of Health."

The reasons for the existence of two separate oaths are no longer very clear.

Although the Office of the Registrar General might prefer the legislature to establish policy on confidentiality and data dissemination under the Vital Statistics Act and does not see itself as having much discretion in these areas under the Act, it seems fairly evident that the Deputy Registrar General has some discretionary power under the statute. It is the policy of the Office "to provide reasonable access to authorized researchers, government representatives and the private sector to vital statistics to enable them to produce more effective results in health, social and economic planning." This policy is implemented with the general understanding that an individual respondent has a right of privacy for personal information compulsorily divulged to the Office. There is no release of identifiable data in general statistical tabulations prepared by the Office, such as the annual volume of vital statistics. The Registrar General does the coding and keypunching of individual

TABLE II: OATH OF SECRECY
OFFICE OF THE REGISTRAR GENERAL

OFFICE OF THE REGISTRAR GENERAL

MACDONALD BLOCK, PARLIAMENT BUILDINGS, TORONTO, ONTARIO M7A 1Y5

Form 30

The Vital Statistics Act

OATH OF SECRECY

I, _____
(given names)

_____
(surname)

solemnly swear that I will hold secret and will not disclose to any person any information given me from the records in the Registrar General's office or obtained from those records by reason of my access thereto except information required in the performance of the duties in my office or information required by a court of law for the purposes of an action, prosecution or other proceeding.

SWORN before me)
)
at the City of)
)
Toronto in the) _____
) (signature of deponent)
Municipality)
)
of Metropolitan Toronto,)
)
this day of 19)

_____
A Commissioner, etc.

data and sends computer tapes to Statistics Canada, which prepares the tabulations. The individual entries on these tapes are not in themselves identifiable, but there is a registration number for each entry that can be linked to the annual index and the original registration. The Registrar General also sends to Statistics Canada an annual identifiable index of birth, marriage, and death registrations. Thus anyone seeking access to vital statistics from Ontario for research and statistical purposes can initially approach Statistics Canada in order to obtain access. Any dissemination of Ontario data by Statistics Canada also has to be authorized by the Registrar General of Ontario.

The Office of the Registrar General does make identifiable data on mortality directly available to researchers. Approximately a dozen epidemiologists were assisted in this fashion during 1977. The conditions for such access are not formally established by written policies. A researcher is normally queried about the sponsorship of his or her work, its value to the general public, the validity and utility of the research, and the extent, if any, of substantive support from the Ontario Ministry of Health. Refusals of access to researchers generally occur because of the lack of supportive documents, the failure to obtain research grants, or a determination that a particular piece of research does not meet the criteria of public good.

It is suggested that the Office of the Registrar General currently perceives less public support for academic research in general. Table Three reproduces a set of standard written undertakings governing access to data, which are imposed on particular research projects. These were formulated by the Office of the Registrar General. The researcher has to supply a credible individual to perform the search of available records. Such an individual is sworn in by means of an oath of secrecy.

Examples of Data Dissemination

The Office of the Registrar General has provided research assistance to a wide variety of users, including such Ontario government organizations as the Occupational Health Branch and the Workmen's Compensation Board of the Ministry of Labour, and the Ministry of Health. Similarly, a wide variety of services have been performed for qualified individual researchers. It is possible, for example, for the Registrar General to furnish a researcher with the causes of death for a specific group of employees who died in Ontario and worked in a specific industry. This type of assistance is of substantial importance in the field of occupational health.

TABLE III: AN EXAMPLE OF TERMS AND CONDITIONS FOR A RESEARCHER SEEKING ACCESS TO ONTARIO VITAL STATISTICS RECORDS FROM THE OFFICE OF THE REGISTRAR GENERAL

a) Provide the manpower and revenue resources to search the records where required;

b) Persons to do the job shall first take an oath of secrecy under the Vital Statistics Act;

c) Guard against the risk of accidentally improper disclosure associated with the proposed method of handling and utilizing the data;

d) Stipulate that in the release of information no identifying data will be given that is restricted by legislation;

e) Outline to me the security arrangements proposed to ensure confidentiality is maintained;

f) Affirm that no contact will be made with the relatives of the deceased;

g) Submit to me, if requested, a draft prior to the release of any report or resultant data obtained as a result of your study;

h) If not contained in your opening correspondence describe the
   i) intended use of the resultant data and
   ii) probable value and significance of results and in particular how the study serves the public interest and good;

i) Provide supportive documents when required;

j) Access to the records is not to be used as a lever to secure funding of the project. Written proof of receipt of grant to be submitted in advance; and

k) Duplicate work is not being undertaken by other researchers.

Source: Mr. N.A. Vetere, Deputy Registrar General, June 8, 1978, Office of the Registrar General, Queen's Park, Toronto.

The Office of the Registrar General has for twenty-five years carried on a death abstract service for the Ontario Cancer Treatment and Research Foundation (OCTRF). The latter receives a card on every patient dying in a particular year where the medical certificate of death mentions cancer. The annual series of cards on every patient are used to produce yearly statistics on cancer. OCTRF uses the identifiable data on individuals to follow-up particular cases for research purposes. Any dissemination of this data for research purposes outside OCTRF would have to be reviewed by an internal committee. The Office of the Registrar General also assists OCTRF in arranging for a large number of death searches to be conducted: "This latter support is invaluable to the regional treatment centers in patient follow-up, and in the conduct of statistical reviews and epidemiological studies."[16]

Another example of specific research assistance by the Office of the Registrar General involved a study of crib deaths. Three of the four authors were associated with the Department of Community Health and Epidemiology of Queen's University.[17]

16 A.H. Sellers, et al., "Statistical Report on Cancer in Ontario, 1975" in Ontario Cancer Treatment and Research Foundation, Cancer in Ontario 1976 (Toronto, 1976), p. 69.

17 A.S. Kraus, R. Steele, M.G. Thompson, and P. deGrosbois, "Further Epidemiologic Observations on Sudden, Unexpected Death in Infancy in Ontario," Canadian Journal of Public Health, LXII (May/June, 1971), 210-219.

This particular study involved interviews with the mothers of eighty infants, who had died suddenly and unexpectedly, and with eighty matched surviving infants, who were used as controls. The Office of the Registrar General for Ontario forwarded to the researchers photocopies of the death certificates for infants who died at age 28 days to 364 days between February 1, 1968 and December 31, 1969, who resided in a specified group of counties, and whose stated causes of death were suggestive of a sudden unexpected death. The counties studied included 16 in south-eastern Ontario and a number in the northeastern region of Ontario. The researchers obtained confirmation of the sudden and unexpected nature of the deaths from physicians signing the death certificates. "After permission was granted by the physician, the parents of the deceased infants were contacted by mail requesting an interview."[18] The research group was able to complete interviews with the parents for eighty cases. The researchers simply state that "an interview was not carried out for three cases."

The assistance of the Office of the Registrar General in the crib death study permitted the important methodological contribution of a matched control group:

18 Ibid., p. 211.

> "The Office of the Registrar General for Ontario located the birth certificates for the SUD infants and then identified the nearest birth certificates in the same files for surviving infants who matched the SUD infants on date of birth (within five days), birthweight (within eight ounces), and area of residence at birth. An interview was completed on one control infant for each of the eighty interviewed SUD cases."[19]

In addition to developing some possible etiologic hypotheses based on the findings of this research project, two major new and independent findings about crib death involved "significant associations between sudden, unexpected deaths in infancy and: (1) mother age 20 or less at the birth of the infant and with one or more prior live births, and (2) a pattern throughout the life of the infant of little time outdoors."[20]

Professor Terence W. Anderson of the Department of Preventive Medicine and Biostatistics of the University of Toronto has been a substantial user of mortality data from the Office of the Registrar General since the mid-1960s. Anderson has produced a number of articles based in part on access to death certificates at the Office of the Registrar General. A study of ischemic heart disease and sudden death concentrated on deaths occuring among men aged 45 to 64 in the province of Ontario between 1901 and 1961.[21]

19 Ibid., p. 211.

20 Ibid., p. 218.

21 Terence W. Anderson and W.H. LeRiche, "Ischaemic Heart Disease and Sudden Death, 1901-61." British Journal of Preventive and Social Medicine, XXIV (1970), 1-9.

The researchers derived information directly from original death certificates.

> "A sample of Ontario death certificates was examined from each of the census years (ending in 1) between 1901 and 1961. The number of certificates examined from each year ranged between 2,500 and 5,000 (Table One), each annual sample being composed of three to five sub-samples drawn from different geographical regions of the province, in which the proportions of urban and rural population were similar to that for the province as a whole."[22]

The researchers recorded diagnostic and other information from those certificates in which the patient was a man aged between 45 and 64 and in which any of the pertinent diagnoses appeared in the certificate. The study concluded that there indeed had been a genuine increase in mortality from ischemic heart disease between 1931 and 1951.

In another article, Anderson and his associates studied the correlation between sudden death from ischemic heart disease and the hardness of local water supply. The study "examined the variation in death rate from ischemic heart disease in three regions of Ontario, according to the hardness of the local water supply and whether or not the death was certified by a coroner."[23]

22 Ibid., p. 2.

23 Terence W. Anderson, W.H. LeRiche, and J.S. MacKay, "Sudden Death and Ischemic Heart Disease: Correlation with Hardness of Local Water Supply," New England Journal of Medicine, CCLXXX (April 10, 1969), 805-807.

In connection with this study, the Office of the Registrar General re-sorted the 55,000 death certificate cards for 1967 according to new categories. This series of articles by Anderson and his associates produced a number of hypotheses which they have proceeded to test by other forms of research, such as autopsies. In some current research into the causes of the higher rates of sudden death from heart disease in northern Ontario, Anderson uses the death certificates to find out the normal place of residence of an individual who has died suddenly in northern Ontario.[24]

In another ongoing study, Anderson is monitoring the mortality patterns of employees of Ontario Hydro because of the risks of radiation to employees in nuclear plants. The research compares the mortality experience of employees of nuclear plants, coal plants, and general Hydro employees. Although most of the data for the study derive from the insurance records of Ontario Hydro, Anderson still has a few occasions each year to search for death certificates at the Office of the Registrar General for employees of Ontario Hydro who have died.

24 See T.W. Anderson, L.C. Neri, G.B. Schreiber, F.D.F. Talbot, and A. Zdrojewski, "Ischemic Heart Disease, Water Hardness and Myocardial Magnesium," Canadian Medical Association Journal, CXIII (August 9, 1975), 199-203.

As a substantial and satisfied user of the services to researchers offered by the Registrar General's Office, Anderson is not aware of any problems concerning access for other researchers. In 1972 Anderson was furnished an opportunity by the then Deputy Registrar General to comment on the final draft version of the new death certificate form. He was able to suggest the importance of various elements of information in the medical certificate of death to the research community. Anderson is also of the opinion that more information from the medical certificate of death and the statement of death should be put into machine-readable form so that it can be more usable for research and statistical purposes. In this connection it would seem desirable for the Office of the Registrar General to set up a small advisory committee on methodological and access issues. There is a comparable problem in the fact that the Office of the Registrar General generally does not allow researchers to contact the next of kin of the deceased or his or her physician. This could create problems for a researcher attempting to undertake retrospective studies. An advisory committee could review proposals for research involving contact with relatives of the deceased. An appropriate research protocol should involve authorization by the physician for the deceased before any re-contact is made with the family. Such an ad hoc arrangement could be approved by the Office of the Registrar General.

The Relationship with Statistics Canada

The Office of the Registrar General indirectly supports significant experiments in the field of record linkages for health research by means of its association with Statistics Canada. Each year the Office sends to Statistics Canada vital statistics on Ontario individuals in the form of computer tapes and alphabetical indexes. The individual data from Ontario are entrusted to the Health Division of Statistics Canada, where they are protected by the strict provisions of the Statistics Act. The secrecy clause of this statute does not permit the release of identifiable data on individuals, except when a particular respondent permits the release.[25] Statistics Canada organizes the vital data from each of the Canadian provinces for production into national statistics. In particular, the agency has organized mortality records from 1951 to the present into a mortality data base or national death index. Statistics Canada has also recently developed the technical capacity to link up its vital data with identifiable individual data brought into the agency by an outside researcher for research and statistical purposes. As a general rule, the agency only releases statistical tabulations from such a research exercise. When a

25 Statistics Act, 19-20 Eliz. II, c. 15 (1971), s. 16; see also David H. Flaherty, Privacy and Government Data Banks. An International Perspective (New York, 1978, forthcoming), chapter 12.

particular research project requires expertise which Statistics Canada does not have, the agency may hire the expertise for the duration of the project. Such talent may be drawn from individuals associated with a particular project. The person or persons hired may have access to individual data as required by the particular project. Under such circumstances, the person or persons are subject to all of the terms and conditions, including penalties, of the Statistics Act.

The issue of record linkage remains a sensitive one at Statistics Canada today, as the agency seeks to formulate a definite policy and code of ethics for record linkage generally.[26] The Statistics Act does not prevent data linkages within Statistics Canada, so long as identifiable results are not released. The agency is naturally concerned about the extent to which the general public will tolerate such an activity. At the Bellagio Conference in August, 1977 Michael Francino of Statistics Canada pointed out that "in some areas, notably health, a good case can be made that the more extensive use of record linkage, coupled with some policy of selected access for bona fide research purposes, would benefit the community as a whole, and more rarely, the individual who might be warned of

26 This paragraph is adapted from ibid., chapter 15.

some condition posing a threat to health."[27] Yet Statistics Canada does not want to appear to be in a situation of disseminating confidential information. It might determine in future that the special needs of medical research should be met through the creation of an organization like a National Center for Medical Research, but this seems unnecessary. Necessary record linkages, especially in the medical field, should take place within Statistics Canada in order to take advantage of its status as a protected data enclave and of the data already collected there. As will be seen below, the linkage of employment records and mortality data has great potential for significant research results. Yet there is no question that record linkage remains a very sensitive public issue, and that Statistics Canada will move forward cautiously in this area.

Statistics Canada regards itself as the custodian and not the owner of the vital statistics obtained from the Office of the Ontario Registrar General and other provincial statistical offices. Statistics Canada does not normally return any identifiable individual data to the Office of the Ontario Registrar General. On occasion the agency has returned historical data. Section 16 of the Statistics Act does permit

27 Michael Francino, "Privacy and Access to Government Microdata for Research and Statistical Purposes, The Canadian Case," (Statistics Canada, Ottawa, July, 1977), p. 3.

Statistics Canada to return data to persons or organizations, such as the Office of the Registrar General, that are respondents originally. But Statistics Canada cannot make data available for administrative purposes without coming under the regulation of the new Canadian Human Rights Act.[28]

Statistics Canada's custody of Ontario vital statistics is also governed by Orders in Council issued by the Governor General in Council in 1919 and 1945 "covering Agreements between the Government of Canada and the Provincial Governments with regard to Vital Records and Statistics."[29] These agreements were the products of conferences in 1918 and 1944 between officials of the federal governments and representatives of provincial vital statistics offices. The 1919 Orders in Council authorized the Dominion Bureau of Statistics, the predecessor of Statistics Canada, to receive from the provinces for purposes of compilation, tabulation and publication on a national basis transcripts or certified copies of the original returns of marriages, births, and deaths from a province. The 1945 Orders in Council authorized the establishment of a Vital Statistics Council for Canada and the creation of a National Register of

28 Canadian Human Rights Act, 25-26 Eliz. II, C. 33 (1977).

29 Privy Council 693, April 22, 1919, and Privy Council 4851, July 31, 1945.

Vital Records. Two general provisions in the Regulations under the 1945 Orders in Council govern uses of the data in question:

> "15. The National Register shall be used by the respective governments for the purposes of verification and statistics only. Any further use shall be approved by each province in respect to its part of the records in the National Register after recommendation by the National Council of Vital Statistics for Canada.
>
> 16. The microfilm copies and any indices compiled therefrom shall be subject to the same restrictions as are imposed by the Secrecy Clauses of the "Statistics Act" of the Dominion, except in such cases as may be determined by the said Council from time to time, and upon authorization from a Provincial Government in respect to its own records."

These general provisions obviously establish the right of the Ontario Registrar General to control uses of Ontario vital statistics in the custody of Statistics Canada for epidemiological purposes.

In practice, Statistics Canada will permit access to Ontario vital statistics for statistical purposes without the explicit permission of the Registrar General, since a researcher will only obtain tabulations. Statistics Canada also tries to persuade researchers that there is no need for them to obtain access directly to identifiable microdata, but that statistical tabulations will be satisfactory. When, as in the case of the Ham Commission, a need for access to identifiable microdata is determined to exist, directions are sought from provincial Registrars General as required. During the last several years the volume of written communications for such purposes between

the Ontario Registrar General and Statistics Canada has been very modest.

From September, 1974 to June, 1976, an Ontario Royal Commission conducted an investigation into matters related to the health and safety of workers in Ontario mines.[30] James M. Ham, now the President of the University of Toronto, was the Commissioner. Professor David Hewitt of the Department of Preventive Medicine and Biostatistics, University of Toronto, conducted two major studies for the Commission on the mortality experience of Ontario uranium miners. The two studies were entitled "The Mortality Experience of Persons on the Uranium Nominal Roll, 1955-74" and "Radiogenic Lung Cancer in Ontario Uranium Miners, 1955-74."[31]

Hewitt began his research on the mortality experience of uranium miners in Ontario by using the uranium nominal roll of the Ontario Workmen's Compensation Board. It is a computer file derived from a card index of all individuals certified to work in an underground mine. The information includes medical data of various types, including x-rays. In maintaining this index the Workmen's Compensation Board automatically links together

30 Report of the Royal Commission on the Health and Safety of Workers in Mines (Toronto, 1976).

31 Ibid., Chapter 3 and Appendix C.

episodes involving the same person. However, the Board is not in a position to collect mortality data on all current and former miners. For this purpose, Hewitt turned to the mortality data base or national death index of Statistics Canada. Since this latter index contains mortality information on all of the Canadian provinces, he could discover the deaths of Ontario uranium miners who may have moved elsewhere in Canada and died there at a later date.

This particular effort at record linkage by the Health Division of Statistics Canada was its first major assistance to an external researcher. The agency produced a series of possible matches between individuals on the nominal roll and registrations of mortality. As directed by the Registrars of Vital Statistics, Statistics Canada made the possible matches available to Hewitt for final resolution. The agency did not have the expertise or the resources available at that time for the work. Thus Hewitt could determine the most likely matches among names. Again as directed by the provincial Registrars of Vital Statistics, Statistics Canada made copies of certain death certificates available to Hewitt for epidemiological studies. The data furnished to Hewitt were in the custody of the Royal Commission and the Workmen's Compensation Board. It is not clear whether or not Hewitt was sworn-in under the Statistics Act in this particular case. It is certain that Statistics Canada obtained

the approval of each provincial Registrar General, including Ontario, before Hewitt could work with identifiable microdata. It is unlikely that Statistics Canada would now permit an outside researcher access to identifiable microdata without swearing the individual under the Statistics Act.

Current attitudes of provincial Registrars General (excluding Ontario, but including Quebec), and perhaps restrictions imposed under interpretations of the federal Statistics Act, could make it difficult to replicate Hewitt's research in future, as the Ham Report recommended. The Occupational Health Branch of the Ministry of Labour in Ontario will continue to monitor the mortality experience of uranium miners, but the staff involved will not enjoy access to identifiable data. This will make it impossible for the research staff to check the accuracy of the specific links and to pursue scientific analysis and linkages with other data, such as the smoking habits of the sample.

The work undertaken by Hewitt for the Ham Commission clearly was a research and statistical use of personal data. But his analysis of the records of particular uranium miners also made the Workmen's Compensation Board aware that a number of persons had died from a particular form of cancer, should have been receiving special pensions, and were not. But using research and statistical data for administrative purposes, even of the

most desirable types, breaches the principle of establishing a strict, functional separation between research/statistical uses of data and regulatory/administrative. Apparently in this particular case, the Registrar Generals of particular provinces, who are the ultimate owners of this data, approved contact with the relatives of the deceased by the Workmen's Compensation Board in order to correct any existing injustices in the award of pensions.

The National Cancer Institute of Canada funds a unit of cancer epidemiology in the Department of Epidemiology and Biostatistics of the University of Toronto. This epidemiology unit is indirectly a major user of mortality data through record linkage projects at Statistics Canada. Under the general direction of Dr. A.B. Miller, the Director of the unit, there are four record linkage projects with Statistics Canada that involve Ontario data. In these cases all of the record linkage work is done at Statistics Canada and only tabulations are released to the researchers. However, the senior statistician for the unit, Dr. G.R. Howe, regularly visits the Health Division of Statistics Canada in Ottawa and participates in the design of the record linkage projects. He has been sworn-in under the Statistics Act.

The first study involves a national registry of tuberculosis cases maintained at Statistics Canada. The agency has linked admission and discharge information on individuals from sanitoria for 1951 to 1960 with national mortality data. There will be an additional linkage for national cancer data from 1969 to 1975. The goal is the evaluation of suspect carcinogenicity in the use of a particular drug for the treatment of tuberculosis. No final results are yet available.

A second study that is underway is based on a five to ten percent sample from the occupation-book renewal file of the Unemployment Insurance Commission for 1965 to 1971. Social insurance numbers for members of the sample are used to link with the unemployment insurance original file data in order to obtain identifiable information. The data are then linked to the national death index at Statistics Canada in order to produce death rates by occupation. The occupation file will also be linked with the cancer data.

No linkages have yet been accomplished for a third study. Researchers for the epidemiology unit collected cases of pulmonary tuberculosis treated in sanitoria between 1930 and 1952. They extracted identifiable data on individuals and their treatment. Forty percent of the sample received treatment in the form of x-ray fluoroscopy. The research will test the

long-term effects of radiation in the form of breast and lung cancer.[32] Linkages with the national death index at Statistics Canada will be performed later in 1978. The project may encounter problems collecting information on deaths occurring before 1950, when the mortality data base of Statistics Canada begins. The provincial Registrars General will have to be approached to obtain the necessary information.

The fourth study is a follow-up examination of the increased risk of bladder cancer among railroad workers. The project has access to data on CNR employees, which will be sent to Statistics Canada and linked with mortality data, to study death rates by the various occupations of workers within the employ of CNR.

A final example of major research indirectly involving Ontario mortality data is a retrospective mortality study of Ontario mickel workers. This is a collaborative venture between the Joint Occupational Health Committee (JOHC) of Inco Metals Company-United Steel Workers of America and the Occupational Health Program at McMaster University. The co-principal investigators are Professor Robin S. Roberts of McMaster University and Dr. Ernest Mastromatteo, the director of occupational health for Inco Limited.

32 See Howard B. Newcombe, Cancer Following Multiple Fluoroscopies, Report No. AECL-5243 (Atomic Energy of Canada Ltd., Chalk River, Ontario, August, 1975).

The data base for the Inco study involves an historical cohort of 50,000 men who have worked for the Ontario division of Inco Metals Company since 1950. Their mortality experience will be ascertained and cause specific mortality rates will be estimated and compared between various occupational groupings of study subjects. The objective is to identify any occupational risks and if possible to isolate the corresponding causal agents. After work history files for each of the individuals in the sample have been prepared by the researchers, Statistics Canada will be asked to do a matching study for mortality experience after 1950. This again raises the problems about confidentiality under the Statistics Act, which were discussed above with reference to other follow-up studies. The Committee on Microdata Release of Statistics Canada has to determine whether or not individual records can be removed from the agency, even if all identifiers are removed; the researchers only need anonymized microdata for analyses. There is also the issue of whether or not the researchers can themselves manually check the sensitive issue of whether or not exact matches have been achieved at the agency by having access to all the relevant identifiable data. In the current case two members of the research team were routinely sworn-in under the Statistics Act. Through Statistics Canada, provincial registrars of vital statistics may also not be willing to furnish the researchers with the death certificates of the particular individuals, so that the details of diagnoses

can be analyzed. The Inco study is using the basic record linkage methodology developed for the studies discussed above.

The proponents of this restrospective mortality study of Ontario nickel workers argue that the successful execution of this study will be of fundamental significance to the nickel industry specifically and to occupational health in general in the years to come. The Joint Occupational Health Committee of Inco and United Steel Workers asserts that "the results of the study will help the Company identify and eliminate any hazardous exposures that may exist and will prove invaluable to governmental authorities who must set guidelines or standards for occupational exposure in the nickel industry. We believe this study is an excellent example of what applied research and occupational health should be - it seeks answers to legitimate questions regarding the health of thousands of Ontario workers."

Recommendations

1. The Office of the Registrar General should continue to provide access to vital statistics for legitimate research and statistical purposes through its own Office and the services of Statistics Canada.

2. The Office of the Registrar General should formulate general rules and regulations for granting controlled access to vital

statistics for legitimate research and statistical purposes. The right of approved researchers to obtain access to such data and the general criteria for access should be set out in the Regulations under the Vital Statistics Act.

3. The Office of the Registrar General should continue to impose written undertakings for the protection of confidentiality on users of identifiable data. These should specify in explicit detail the research which is to be undertaken and the uses to which the data can be put.

4. The Office of the Registrar General should establish an advisory committee to assist in the formulation of policies on confidentiality and the dissemination of data and to review specific research proposals from a scientific and ethical perspective. A small advisory committee of this type should include representatives from the general public, the Office of the Registrar General, and the research community. Although the Deputy Registrar General has the primary responsibility to protect the privacy of individuals whose data are in his custody, this advisory committee can assist in the essential task of balancing personal privacy against other important societal interests. Even though an avalanche of researchers seems most unlikely to descend upon the Office of the Registrar General, it should be recognized that the Office can only accommodate and service a limited number of research and statistical users. If these resources become too constrained, the advisory committee could assist in attempting to devise needed solutions.

5. The penalties for breach of secrecy under Section 48 of the Vital Statistics Act should be substantially increased. Specific sanctions should be attached to breaches of the oath of secrecy prescribed under Regulation 820.

CHAPTER II

CENTRAL STATISTICAL SERVICES

Central Statistical Services (CSS) is a program of the Ministry of Treasury, Economics, and Intergovernmental Affairs. In succinct words this program provides a central statistical service to the government of Ontario by administering the Ontario Statistics Act; maintaining liaison with Statistics Canada; collecting, processing and disseminating statistics; supplying statistical advisory services; and providing economic accounts data and social statistics, including demographic studies. CSS has an annual budget of 1.5 million dollars and a staff of about 45 persons. The staff are organized into four basic sections: Liaison and Client Services; Data Management; Surveys and Advisory Services; and Social and Economic Data. Table IV is a listing of publications by Central Statistical Services as of March, 1978. All of the data are aggregated in character.

In actual fact CSS has a limited range of ongoing responsibilities especially in the area of personal data, because of statistical activities undertaken by individual ministries. Although they are a body of resources that other ministries of the Ontario government call upon for assistance in statistical methodology,

ad hoc surveys, and information services, ministries do a lot of their own work. There are substantial statistical resources, for example, located in the Ministries of Labour, Health, and Transportation and Communications. CSS does attempt to supplement the data already made available for Ontario by Statistics Canada. But in fact Ontario has a highly decentralized statistical system, and no single statistical bureau as such. One of the current goals of CSS is never to duplicate the activities of other government agencies.

In addition to the provision of statistics and the design of surveys, CSS has an advisory, coordinating, and liaison function. It has the right to advise the Management Board of Cabinet concerning statistical policy, but coordination is a low-profile activity. CSS also supports a Statistical Data Users' Committee with representatives appointed from each ministry. This Committee is an advisory body to CSS and a semi-annual forum for discussions between CSS and the various Ministries. Its terms of reference suggest that it will provide a means whereby CSS "may ensure that all Ministries are informed of significant developments in the field of government statistics."

The generally-limited functions and responsibilites of CSS, including the maintenance of a low-profile and the lack of control functions, were re-affirmed by a report of the Ontario

TABLE IV: PUBLICATIONS
CENTRAL STATISTICAL SERVICES

Demographic:

Short-Term Population Projections, 1975-1986, Dec. 1976
Ontario Population Trends: A Review of Implications, Dec. 1976
Ontario Population Estimates, By Planning Regions and Counties June 1972 - June 1976, Dec. 1977
Monthly Demographic Bulletin

Economic Data:

Ontario Economic Accounts - Quarterly Time Series 1947-1975, Sept. 1977
Ontario Economic Accounts - Quarterly Bulletin
Ontario Statistics 1977, 2 Volumes
Credit Union Quarterly Statistical Bulletin
Consumption of Fuel and Electricity by Ontario Manufacturing Industries, 1974 data

Social Data:

Social Indicators for Ontario

Government and Miscellaneous:

Index of Ontario Government Statistics for Municipalities
ONSTAT News (C.S.S. Quarterly)
Index of Statistical Files in the Ontario Government, 1977

Computerized Files - Access Through Central Statistical Services:

Credit Union Statistics
Census of Industries, Statistics Canada, Ontario Data
Census of Population, Statistics Canada, Ontario Data
Population Projections Ontario and Counties 1971-2001

Source: Central Statistical Services, Demographic Bulletin (March, 1978), p. 11.

Interministerial Committee on Provincial Statistics at the end of 1977. These conclusions were subsequently approved and summarized in a directive on "Statistical Services" issued by the Management Board of Cabinet on April 18, 1978. The directive determined the scope of statistical services and instituted a cost recovery approach between CSS and its clients in the ministries.

Central Statistical Services cannot be described as a statistical agency in comparison either with other Canadian provinces or the state statistical agencies in other federal systems, such as West Germany. Quebec, for example, has a substantial statistical bureau with a staff of 250. Ontario used to have an Ontario Statistical Centre, but this has been changed to "Services." The Ontario Committee on Government Productivity in 1973 did recommend the creation of a central Ontario statistical bureau. In fact, Central Statistical Services has had a tenuous existence, was almost disbanded in 1978, and has recently been reorganized. It has been reduced in both personnel and duties; for example the mining and logging census has reverted to the Ministry of Natural Resources. The financial resources of CSS have been reduced by one-third and cost recovery introduced.

One rationale for the current situation in Ontario is as follows. The province relies heavily on Statistics Canada, since 40 per cent of the data collected by the national agency pertains

to Ontario. Provincial ministries also like to do business directly with Statistics Canada. For legal reasons that will be discussed further below, CSS also does not have any particular advantages in terms of dealing with Statistics Canada. Secondly, there is no strong government orientation to central statistical services as such; ministries want to do their own statistical work. It is further suggested that the government is not heavily dependent on research and statistical data in the process of decision-making; there is a low priority for research and statistics. An aura of suspicion surrounds this field. The politicians and policy-makers are also worried about the burden on householders and small businesses of collecting statistical information. Thus the situation has arisen where other ministries only use CSS when they have to and when they can afford to pay for the services. It is also alleged that there is a tendency to avoid CSS because it is located within the powerful Ministry of Treasury, Economics, and Intergovernmental Affairs.

Data Collection

CSS essentially does not collect personal data on an ongoing basis. It does not conduct any regular surveys but carries out a number of ad hoc ones for various ministries. An example of a

recent ad hoc survey conducted under the auspices of the Ontario Statistics Act was the health survey of Parry Sound and vicinity for the Ministry of Health. This ministry is studying the needs of persons who are at or approaching retirement age in a particular area. The study focuses on the health of these citizens and the various health services that they require. CSS conducted an interview survey with a cross section of persons living in the Parry Sound area. They are now doing similar surveys for Kenora and Rainy River.

CSS also produces interprovincial migration estimates on the basis of administrative data obtained from the Ontario Motor Vehicle Bureau and data about federal family allowances. CSS obtains addresses but not names from the "change of address of driver license holders" file. They also obtain data from a file known as the "transfer of family allowance accounts."[1] CSS began to study this topic because the provincial Ministries of Health and Housing were interested in it. Third ministries are willing to furnish data of this type to CSS.

CSS also works on estimating population and labour force for small areas in Ontario. They use microdata from the assessment files of the Ministry of Revenue, but without names. They obtain

1 Central Statistical Services, Demographic Bulletin (March, 1978), 2.

vital statistics from the Office of the Registrar General and Statistics Canada without names, addresses, or registration numbers of individuals. Some of this work is for the Ontario Ministry of Labour. CSS has no access to federal or provincial tax data on an individual basis. It obtains no individual data from Statistics Canada, nor does it make use of the recent innovation of public use samples from surveys and censuses made available by Statistics Canada.

The Ontario Statistics Act

Central Statistical Services originally prepared the draft of the Ontario Statistics Act in the early 1960s; it currently administers the Act, which is available to other Ministries, although there is no obligation on them to use it. The Act has the limitation of not naming a single statutory statistical bureau for the provinces. It simply authorizes the Minister of any department of government:

> "a) To enter into an agreement with the Government of Canada, or the government of any province in Canada, or any agency of any such government to provide for an exchange or joint collection of statistical information;
> b) To collect, compile, analyze and publish statistical information;
> c) To collect statistical information jointly with the Minister of any other department of government."[2]

2 Ontario Statistics Act, S.O., 1970, c. 443, s. 2(1).

Surveys undertaken under the authority of the Act are mandatory in character; there are penalties for refusal to answer.[3]

The Ontario Statistics Act includes an oath of secrecy before any "person shall collect, compile, analyze or publish statistical information" under this Act:

> "I, ..., do swear that I will faithfully discharge my duties under the Statistics Act and, except as I may be legally required, I will not disclose or give to any person any information or document that comes to my knowledge or possession by reason of my duties under the Statistics Act. So help me God."[4]

Another section prohibits unauthorized disclosures:

> "4(2) Subject to Section 6, no public servant having knowledge of the answers to questions asked in a questionnaire under this Act shall disclose or give to any person any information or document with respect to such answers without the written permission of his Minister, and, except where statistical information is collected jointly under this Act, such permission shall be limited to the disclosing or giving of information or documents to public servants in the Minister's department or in prosecutions instituted for offences against this Act."

The Act does ultimately provide for the confidentiality of all statistical information and its use only for statistical purposes: "Notwithstanding anything in this Act, no Minister or public servant shall, in any way, use the answers to questions asked in a questionnaire under this Act for any purpose other than the purposes of this Act."[5]

3 S.O., 1970, c. 443, sect. 3, s. 7.

4 S.O., 1970, c. 443, s. 4(1).

5 S.O., 1970, c. 443, s. 4(3).

Section 6 controls the disclosure of information to another department. The contents of a questionnaire collected by one department can only be disclosed to another government department, "where a person who has answered a question in a questionnaire consents in writing."[6] This provision does not apply to an index or list, summary of statistics, or other publication under this Act, that discloses the names or locations of individual firms or businesses, or the types of products commercially produced, manufactured, or dealt with by individual firms or businesses.[7] Finally, the Statistics Act includes criminal penalties for unauthorized acquisition or disclosure of information.[8]

Although the Ontario Statistics Act appears at first reading to be a strong statute from the point of view of confidentiality, it has a variety of significant limitations. The major restrictions concern the level of access to data at Statistics Canada. The federal Statistics Act has in general a very strict provision prohibiting the release of identifiable individual data outside the agency. A major exception to this prohibition is that Statistics Canada may release "the particulars of any

6 S.O., 1970, c. 443, s. 6(1).

7 S.O., 1970, c. 443, s. 6(2).

8 S.O., 1970, c. 443. s. 8.

information obtained in the course of administering this Act" to a statistical agency of a province pursuant to an agreement under Section 10 of the Act.[9] A similar exception to the duty of secrecy exists under Sect. 11 concerning the particulars of any information collected jointly or with a department or a corporation. Because of a determination by Statistics Canada that the Ontario Statistics Act is unsatisfactory, Ontario has to rely on Sect. 11 and cannot qualify under Sect. 10. This means that provincial statistical agencies in provinces such as Quebec, Manitoba, and Saskatchewan, can under Sect. 10 receive information collected in a survey by Statistics Canada, because the Statistics Acts of these provinces include provisions for confidentiality and penalties for disclosure of information similar to the federal Statistics Act.

Section 10 provides a provincial statistical agency with the identifiable replies to specific statistical inquiries. But Section 10 also specifies the necessary characteristics of such a provincial statistical agency. In the first place it seems obvious that Ontario does not have a statistical agency as such. It simply has a Statistics Act that permits a fairly open flow of data among government departments under somewhat controlled conditions. The situation does seem inadequate from the

9 Statistics Act, 19-20 Eliz. II, 1971, c. 15, as amended in 1977, s. 16(1).

perspective of providing strong protections for the confidentiality of statistical data. The second requirement under Section 10 is that the statistical agency of the province must be "prohibited by law from disclosing any information of a kind that Statistics Canada, its officers and employees would be prohibited from disclosing under Section 16 if the information were furnished to Statistics Canada."[10] The secrecy requirements under the Ontario Statistics Act are simply not as strict as the federal legislation.

Thus Ontario has to depend on agreements under Section 11 in order to obtain data from Statistics Canada. This is somewhat less liberal and definitely more cumbersome to implement than Section 10. Section 11 permits Statistics Canada to enter into an agreement with any department or any municipal or other corporation in the province of Ontario "for the exchange of information collected jointly with such department or corporation from a respondent."[11] The respondent has to be informed that the information is being collected jointly on behalf of Statistics Canada and the department or corporation. More importantly, any respondent may object in writing "to the sharing of the information between Statistics Canada and the department or corporation."[12]

10 S.C., 1971, c. 15, s. 10(2)(b).

11 S.C., 1971, c. 15, s. 11(1)

12 S.C., 1971, c. 15, s. 11(2).

Thus under this section respondents, who still must respond to Statistics Canada, are given more rights than usual in the collection of statistical data. This provision can cause problems in determining the statistical reliability of the data in question at least at the provincial level.

There is a risk of exaggerating the distinction between Section 10 and 11 agreements, and thus using legal niceties as a whipping boy for any perceived inadequacies of Central Statistical Services. In terms of access to data for the province, there is little that can be done under Section 10 that cannot also be accomplished under Section 11. For example, few Ontario respondents exercise their rights to refuse data to Central Statistical Services under the Section 11 agreement by which the latter obtains the federal Census of Manufacturing. Thus Section 10 access should not be regarded as the panacea needed to revitalize CSS. The latter's prime utility to various provincial ministries has to involve coordination and the provision of competent technical services for these users. For example, it takes Statistics Canada a substantial amount of time, perhaps two years, to produce tabulations from the Census of Manufacturing. The Alberta statistical bureau is able to receive the raw data from Statistics Canada and produce data relevant to Alberta within a quarter of the time required by the national agency. The tabulations produced quickly by the Alberta

statistical bureau are the type of assistance that Ontario provincial ministries need as well.

CSS alleges that even under Section 10 arrangements with Statistics Canada, it would be subject to too many limitations on further re-dissemination of data to provincial ministries. This is both an accurate statement and a reflection of a desirable state of affairs. Sensitive identifiable data should only be entrusted to a statistical bureau such as CSS under authorized sharing arrangements with the national agency that include detailed specifications about data protection and dissemination. Section 10 and 11 agreements with the provinces are and should be a legitimate vehicle whereby Statistics Canada can share its responsibilities to protect confidentiality with an agency such as CSS. The latter should provide a useful service to other ministries by disseminating only tabulations and anonymous microdata. The only exception from the point of view of confidentiality should be a situation where a Section 10 or 11 agreement binds several parties within a single data enclave for the exchange of data. This could include, for example, Statistics Canada, the Ministry of Colleges and Universities, and Central Statistical Services. The allegation that statistical bureaus should not become statistical cemeteries is a valid one, except that concern for confidentiality should also prevent a statistical bureau from becoming a leaky dike. The general public should be discouraged from entrusting

any type of personal data to a government organization, especially a statistical bureau, that does not maintain the highest standards of confidentiality. One of the strongest justifications for the existence of CSS should be its broker role with respect to the re-dissemination of Ontario data in a manner suited to the needs of separate Ontario ministries.

The current Ontario statistical situation can be usefully contrasted with the current situation in British Columbia, which passed a new Statistics Act in September, 1977. The explicit purpose of the legislation is to create a statistics agency, under the authority of a director, with provisions on data gathering and protection of privacy similar to Statistics Canada. The explanatory notes to the legislation specifically indicate that improved access to data from Statistics Canada will be a benefit of the new legislation. The Statistics Division in British Columbia is in fact a part of the Ministry of Economic Development. It is a new agency that is intended to be small in character, but it does have the benefit of strict legislation on secrecy. It does not intend to collect data directly from respondents but to make better use of Statistics Canada data and administrative data already in existence. Like Ontario, it will do ad hoc surveys for other provincial ministries. The specific provisions of the British Columbia Statistics Act are in fact very similar to the Ontario legislation. However, the British

Columbia legislation essentially reinacts the provisions on secrecy under the federal Statistics Act, including Section 10 and Section 11.[13]

The fact that the Ontario government ministries cannot obtain access to data from Statistics Canada under Section 10 is an unnecessary limitation. The basic need is for the creation of a statutory statistical agency in Ontario such as British Columbia has recently instituted. CSS does want access to data transfers under Section 10. But attempts at a new provincial Statistics Act for this purpose have been deferred by the feeling that to do so would be to encourage more questionnaires to households and small businesses.

Confidentiality in Data Dissemination

To the extent that CSS is able to function as a statistical agency, it maintains the usual traditions of concern for confidentiality in data dissemination. In its ad hoc health survey of Parry Sound and vicinity for the Ministry of Health, respondents were informed by letter that "the survey is being conducted on a strictly confidential basis. The answers you

13 British Columbia Statistics Act, S.B.C., 1977, c. 63, s. 9, 11, 12.

give will be used only to generate aggregate statistics relevant to the study; names of individuals will not be divulged." Some months after this survey was completed, the Information Systems Division of MOH authorized CSS to destroy the individual questionnaires, "since this information is of a confidential nature." The provisions of the Ontario Statistics Act discussed above do require the protection of the confidentiality of statistical data, but also permit fairly widespread dissemination within the provincial government.

The Interministerial Statistics Committee currently has in existence a Working Group on Confidentiality and Availability of Information. It is a committee addressing technical issues and not general policy. The Working Group was created because the larger Committee was "concerned about increasing the availability and utility of information within the Ontario Government while at the same time maintaining required levels of confidentiality. This concern has become more acute in recent months, since planned retrenchment by Statistics Canada will require increased reliance on our own resources."[14] The Working Group will examine the technical and operational requirements to ensure necessary levels of confidentiality while providing for the maximum availability of information. Given the difficulties

14 Memorandum, E.P. McCoy, Central Statistical Services, May 4, 1978.

discussed above, it seems unlikely that this Working Group will be able to address the fundamental difficulties hindering access to statistical data within Ontario.

Recommendations

In light of the discussion above, it should be obvious that Ontario is paying an unnecessary price in terms of access to data by not having a provincial statistical agency as such. It seems evident from the point of view of protecting the confidentiality of personal data collected or transferred for statistical purposes, and in order to promote their dissemination for research and statistical purposes, that Ontario should have a statistical bureau along the lines of the Quebec and British Columbia models. This will require a revision of the Ontario Statistics Act that would establish an Ontario Statistical Bureau, create statutory protections for confidentiality at a level comparable to Statistics Canada, and increase the sanctions for breach of confidentiality. A main purpose of the proposed Ontario statistical bureau would be to enter into Section 10 agreements with Statistics Canada that would be of substantial benefit to various provincial ministries in Ontario. This type of development would correspond with the present thinking of Peter Kirkham, the Chief Statistician of

Statistics Canada, about the need to decentralize statistical activities in Canada over the course of the next decade.[15]

The above recommendation was formulated before the author of this Working Paper had an opportunity to review the Report to the Executive Council of the Government of Ontario in March, 1973 by the Committee on Government Productivity, which was made up of high-level business executives and civil servants.[16] To state the matter briefly, the 1973 Report reached the same conclusions and made the same recommendations that have been arrived at independently in the current instance. The Committee found that ministries tried to meet their own statistical requirements and that there was a resulting need to minimize duplication and to exploit government statistics to the maximum.[17] After reviewing a series of options for restructuring, and recognizing the existence of a decentralized system, the Committee advocated "a central statistical agency with decentralized research groups."[18] The functions and scope of activities of the proposed new Ontario Statistical Bureau were outlined in detail, including a strong need for confidentiality:

15 See David H. Flaherty, Privacy and Government Data Banks. An International Perspective (New York, forthcoming, 1978), chapter 15.

16 Committee on Government Productivity, Report to the Executive Council of the Government of Ontario No. 9 (March, 1973), 98-112.

17 Ibid., p. 100.

18 Ibid., pp. 103-105.

> "It would be one of the O.S.B.'s responsibilities to formulate and enforce standards for protecting the confidentiality of data collected for statistical purposes. Special attention should be given to the proper discharge of this function, since the reliability of data collected would rapidly be eroded if the respondents lost confidence in the protection provided. It should be noted, moreover, that certain data are obtainable from Statistics Canada only if confidentiality and security standards are strictly enforced and policed."[19]

It is not self-evident why these 1973 recommendations were not implemented; they remain as appropriate now as when they were first written.

The main reason for the revitalization of CSS should be to better serve the needs of the province for access to data and statistical expertise. CSS simply has too low a priority at present. From the point of view of Statistics Canada, a vacuum exists in the Ontario provincial statistical scene - who speaks for Ontario on statistical matters is a common federal lament. CSS requires more active authority and influence. The Central Statistical Office in Great Britain, which is a comparable coordinating operation, derives its power from considerable influence over the filling of jobs in a highly decentralized statistical system. The Office has real powers of assignment and promotion. It also has considerable power by being directly associated with the Cabinet Office and thus the Prime Minister, and by being headed by Professor Sir Claus Moser, a statistician of considerable personal eminence. Surely there are some lessons here for Ontario. CSS has to take an active role in preventing

19 Ibid., p. 106.

duplication and waste at the provincial statistical level. This can only come about by strengthening its coordinating role.

One method of improving access to data from Statistics Canada that has not yet been adequately explored by CSS is utilization of the new public use samples from censuses and social surveys. This is especially ironic since the Ontario Statistical Centre was one of the original supplicants for such a service in the mid-1960s. Although CSS has purchased the tapes on individuals, households, and families based on the 1971 Canadian census, the availability of data only at the provincial and Metropolitan Toronto levels of geographical identification has resulted in limited interest in the available data on the part of CSS and its clients, who have a strong concern with small-area data. If this is indeed the case, CSS should pursue the development of Regional Statistical Data Bases or statistical data banks, such as are now in operation in West Germany and Sweden.[20] On the other hand, the provincial ministries may yet be unaware of potential applications of public use samples from the census, the Labour Force Survey, and the Survey of Consumer Finance. Significant uses of such public use samples are occurring at the state level in the United States, especially when tapes from a

20 Flaherty, Privacy and Government Data Banks, chs. 7, 10.

national survey are produced for particular states; perhaps CSS could encourage the production by Statistics Canada of public use samples for Ontario only.

A reinvigorated Ontario statistical bureau should also have the capacity to ensure more functional separation in law and practice between statistical and administrative data on individuals in the hands of various Ontario ministries. Public tolerance of the collection of large amounts of personal data for statistical purposes will be increased, if the data are housed in a statistical enclave with strict provisions on confidentiality and data dissemination. In this connection CSS should promote government-wide standards for the protection of confidentiality in data collection and data dissemination in statistical form. The American federal statistical system has recently completed a model effort to promote government-wide standards on the avoidance of disclosure in statistical activities.[21]

21 See Federal Committee on Statistical Methodology, Subcommittee on Disclosure-Avoidance Techniques, Report on Statistical Disclosure and Disclosure-Avoidance Techniques. Statistical Policy Working Paper 2 (Office of Federal Statistical Policy and Standards, U.S. Department of Commerce, Washington, D.C., May, 1978).

CHAPTER III

MINISTRY OF HEALTH

Organization

Since this study is concerned with research and statistical uses of health-related personal data held or acquired by the Ministry of Health, this discussion will focus particularly on the Data Development and Evaluation Branch (DDEB) of the Information Systems Division. This Branch and Division fall within the responsibility of the Assistant Deputy Minister of Health for Administration and Health Insurance. The Information Systems Division, which is the information resource for the Ministry of Health and the health agencies in the community, is headed by an executive director. The two branches of the Division essentially perform a data base function for the Ministry of Health. The Systems Management and Co-ordination Branch "is responsible for the development, maintenance and operation of the computer systems necessary for the operation and administration of the Ministry programs," especially the Ontario Health Insurance Plan (OHIP).[1]

1 Ontario Ministry of Health, Annual Report, 1976/77 (Toronto, 1977), p. 6.

The Data Development and Evaluation Branch "provides access to the data resources of the Ministry to all levels of the health care sector engaged in the planning and evaluation process."[2]

The primary function of the Ministry of Health is to provide a health care delivery system to meet the health needs of the population. Most services of medical practitioners and specified services of other health practitioners provided in the ambulatory or institutional setting are paid for through the Ontario Health Insurance Plan, a ministry-administered plan which covers virtually all Ontario residents. Data are collected essentially for administrative purposes, but also provide significant support for research both within and without the ministry. In its submission to the Commission on Freedom of Information and Individual Privacy in September, 1977, the ministry stated that in the field of research it "is one of the major providers of data from which aggregate information can be obtained to assist in all aspects of research of a short, medium and long term in the medically-related sciences, applied research and operations research."

The Information Systems Division was formed in 1973 by grouping together data processing, systems and programming, and hospital

2 Ibid., p.6.

statistics branches of the Ministry of Health. The addition of an analytical group producing operational analysis and evaluation techniques in 1974 led to the realignment of the Division into the two current branches. The Data Development and Evaluation Branch (DDEB), which is headed by a director, currently employs a total of 59 persons, including 23 economists and statisticians. DDEB produces health care statistics, assists a wide variety of users in identifying the sources and availability of health-related data, and provides analytical evaluation and monitoring reports to the ministry and other agencies and users. Thus DDEB is responsible for publications such as Hospital Statistics 1976 and OHIP Practitioner Care Statistics 1976-77. It can tabulate OHIP claims data for internal and external use, and can also carry out research on health care for the Deputy-Minister of Health on an ad hoc basis. It is worth recognizing that DDEB has both administrative and statistical responsibilities in connection with health-related data on individuals. The sections of the Branch combine several functions with respect to administrative uses of data, research and statistical uses of data, and the preparation of data for various outside users.

Sources of Data

The Ministry of Health obviously accumulates a vast amount of health-related data on individuals in the course of carrying out its varying responsibilities. The data are organized in a series of separate information systems produced by varying administrative activities. The Ministry of Government Services operates three computer centres, one being located on the premises at 15 Overlea Blvd. in the Ministry of Health Building. This computer centre has the physical custody of most Ministry of Health data in EDP form. Approximately 80 percent of the work of this particular data centre is health-related. The Ministry of Health (MOH) specifies the required level of security for each of its files. The data centre is also directed by the person who used to operate it for MOH directly.

No adequate inventory or catalogue seems to exist for the personal data files or bases maintained by the Ministry of Health, although the Information Systems Division has had long-standing plans to produce one. It is thus impossible to prepare one for purposes of this Working Paper. It is clear that the three main building blocks or large files of health information in Ontario are the master file of the medically-insured population, the data on physician services to individuals, and the data on hospital services to individuals.

The DDEB has compiled an inventory of data available for District Health Council Planning for the Ontario Council of Health's Work Group on Data Requirements for District Health Council Planning. It is scheduled for publication in the fall of 1978. The data in the inventory are available through MOH. The compilation lists the sources responsible for the maintenance of the data, points of contact, brief descriptions of the data, methods of storage, and possible delivery times. Data for District Health Councils will be produced in aggregate form only in order to protect personal privacy. The inventory itself in its current format will be useful to outside researchers generally and should also serve as a model for a catalogue of MOH data bases.

The Catalogue of Statistical Files in the Ontario Government, 1977, which is produced "for government use only," lists a substantial number of data files held by the two branches of the Information Systems Division.[3] The Catalogue does not clearly indicate whether the files in question are identifiable individual data or statistical (i.e. non-identifiable) data. One of the largest files is the OHIP Claims File, which is held by the Systems Management and Co-ordination Branch. The OHIP

3 Ministry of Treasury, Economics and Intergovernmental Affairs, Catalogue of Statistical Files in the Ontario Government, 1977 (Toronto, 1977), pp. HL19-HL29. Hereafter cited as Catalogue.

Medical Claims File accumulates records of some 5 million claims per month, which are used to record data for accounting purposes in the administration of OHIP.

The Catalogue also lists a variety of data files held by the Insurance Claims Branch of the Health Insurance Division, the Community Health Protection Branch of the Community Health Division, and the Consulting Services Branch of the Personal Health Division. The variety of files held by this particular group of branches includes hospital discharge reports, data on venereal disease, tuberculosis control, and communicable diseases, and data on maternal mortality and still birth and infant deaths.

The MOH now receives hospitalization data in the form of the Ontario Data File on every patient from the Hospital Medical Records Institute (HMRI). The Ontario Data File "means the totality of data contained in the completed abstracts submitted to HMRI by all hospitals in respect of patients discharged since January 1, 1975 from such hospitals."[4] As received by MOH, this data "shall also contain complete patient identification as

4 Contract between HMRI and the Minister of Health for Ontario, April 1, 1978, s. 1.1(c). Hereafter cited as Contract.

supplied by the hospital in respect of OHIP number, birthdate and sex, together with any unique individual identifier such as the Ministry may require hospitals to supply in the future."[5] MOH uses the Ontario Data File for both statistical and administrative purposes.

An outline of the contents of each confidential file held by DDEB on March 23, 1978 lists a total of 22 data files of varying sorts. Almost all of them appear to be ongoing in character. The specific data files are described as follows:

	Number
Patient: Source Documents	4
Patient: Listings	3
Patient Records - Ad Hoc	
Internal	6
External	1
Physician-Practitioner Records	
Ongoing	2
Ad Hoc	2
Pharmacy Records	2
Lab Listings	2

The available tabulation lists the parent file or source for each of these 22 data files. About one-third are derived from the OHIP Medical Claims File. Other specified patient files or sources include the following: chronic hospitals, psychiatric hospitals and units, HMRI files, collision reports, physician profile system, and Ontario drug benefit file.

5 Contract, Attachment C.

The above information about DDEB simply supports the view that there are a large number of health-related data files within MOH that can and should be used for scientific purposes. The actual number of minor files derived from the three large files is unknown. Each new project or program either for an internal or external user augments the existing data base. One MOH listing of project descriptions named 94 separate data files. Derivative files grow as the government's policies develop. For example, the planning needs of District Health Councils require the development of capacity to produce tabulations of different sorts for a large number of geographic areas in the province.

The Health Division of Statistics Canada also holds identifiable health data on Ontario individuals. The types of data include hospital admissions and separations, the admission and separation of mental hospital in-patients, abortions, tuberculosis, and renal failure. Only the last two data files are said to contain the names of specific patients. There is no identification of individuals in the abortion data. The general data on hospital admissions and hospital separations do not contain names; OHIP numbers are scrambled. In general, Statistics Canada cannot release identifiable individual data under the Statistics Act.

Statutory Provisions on Confidentiality and Data Dissemination

Major statutory provisions regulate the protection and dissemination of personal data held by the Ministry of Health and used for research and statistical purposes.[6] The most relevant statutes are the Ministry of Health Act, the Health Insurance Act, and the Public Hospitals Act.

The Ministry of Health Act grants broad powers to the Minister to promote research and to collect and publish information and statistics on health:

> "6.(2) The Minister in exercising his powers and carrying out his duties and functions under this Act, ...
>
> (c) may initiate, promote, conduct and maintain surveys, scientific and administrative research programs and planning studies into matters relating to the health needs of Ontario and obtain statistics for purposes of the Ministry;
>
> (d) may collect such information and statistics respecting the state of health of members of the public, health resources, facilities and services and any other matters relating to the health needs or conditions affecting the public as are considered necessary or advisable, and publish any information so collected; ...."[7]

6 T.G. Brown of the Commission on Freedom of Information and Individual Privacy, and Barbara Casson Robin of the Royal Commission on the Confidentiality of Health Records generously permitted access to their compilations of relevant Ontario statutes for purposes of this Working Paper.

7 Ministry of Health Act, S.O., 1972, c. 92, s. 6(2)(c)(d).

The Health Insurance Act of 1972, as amended through 1975, governs the operations of the OHIP system. The Act specifies that the Minister of Health may "authorize surveys and research programs and obtain statistics for purposes related to the plan."[8] The Act requires every physician and practitioner who performs an insured service for an insured person to provide the general manager of the Plan "with the particulars of his services and account that are required by this Act and the regulations or the General Manager for the purpose of payment of the claim."[9]

> "S. 33(2) Every insured person shall be deemed to have authorized his physician or practitioner who performed insured services to provide the General Manager with such information respecting the insured services performed as the General Manager requires for the purposes of the Plan.
> (3) No action lies against a physician, practitioner, hospital or related health facility providing insured services or any member of his or its staff because of the furnishing to the General Manager information relating to insured services provided by him or it."

It is arguable that the data so furnished can only be used for administrative purposes and statistics, but not for research.

A major section on confidentiality under the Health Insurance Act is reproduced as Table V. In addition to establishing standards of secrecy of personal data collected in the course of

8 Health Insurance Act, S.O., 1972, c. 91, as amended by S.O., 1974, c. 60 and c. 86, and S.O., 1975, c. 52, s. 2(2)(e).

9 S.O., 1972, c. 91, s. 33(1).

TABLE V: PROVISIONS ON CONFIDENTIALITY IN THE ONTARIO HEALTH INSURANCE ACT

44. (1) Each member of the Medical Review Committee, every practitioner review committee, the Medical Eligibility Committee and the Appeal Board and each employee thereof, the General Manager and each person engaged in the administration of this Act and the regulations shall preserve secrecy with respect to all matters that come to his knowledge in the course of his employment or duties pertaining to insured persons and any insured services rendered and the payments made therefor, and shall not communicate any such matters to any other person except as otherwise provided in this Act.

(2) A person referred to in subsection 1 may furnish information pertaining to the date or dates on which insured services were provided and for whom, the name and address of the hospital and health facility or person who provided the services, the amounts paid or payable by the Plan for such services and the hospital, health facility or person to whom the money was paid or is payable, but such information shall be furnished only,

(a) in connection with the administration of this Act, The Medical Act, The Public Hospitals Act, The Ambulance Act or the Hospital Insurance and Diagnostic Services Act (Canada), the Medical Care Act (Canada), or the Criminal Code (Canada), or regulations made thereunder;

(b) in proceedings under this Act or the regulations;

(c) to the person who provided the service, his solicitor or personal representative, the executor, administrator or committee of his estate, his trustee in bankruptcy or other legal representative;

(d) to the person who received the services, his solicitor, personal representative or guardian, the committee or guardian of his estate or other legal representative of that person; or

(e) pursuant to a subpoena by a court of competent jurisdiction.

(3) The information referred to in subsection 1 may be published by the Ministry of Health in statistical form if the individual names and identities of persons who received insured services are not thereby revealed.

(4) The General Manager may communicate information of the kind referred to in subsection 2 and any other information pertaining to the nature of the insured services provided and any diagnosis given by the person who provided the services to the statutory body governing the profession or to a professional association of which he is a member.

Source: Health Insurance Act, 1972 S.O. 1972, Chap. 91, s. 44 (44(1) as amended by 1974, c. 60, s. 9; 44(2) as re-enacted by 1974, c. 86, s. 2).

administering the OHIP system, section 44 provides that the information collected "may be published by the Ministry of Health in statistical form if the individual names and identities of persons who received insured services are not thereby revealed."[10] The general penalty clause in the Act, which would apply to a breach of secrecy, provides that "every person who contravenes any provision of this Act or the regulations for which no penalty is specifically provided is guilty of an offence and on summary conviction is liable to a fine of not more than $2,000."[11]

The Public Hospitals Act regulates the activities of hospitals in Ontario.[12] It provides that "the medical record compiled in a hospital for a patient or an out-patient is the property of the hospital and shall be kept in the custody of the administrator."[13] Subject to the approval of the Lieutenant Governor in Council, the Minister of Health may make such regulations with respect to hospitals as are considered necessary for "the records, books, accounting systems, audits,

10 S.O., 1972, c. 91, s. 44(3).

11 S.O., 1972, c. 91, s. 50.

12 Public Hospitals Act, S.O., 1970, c. 378, as amended by S.O., 1972, c. 90.

13 S.O., 1970, c. 378, s. 11.

reports and returns to be made and kept by hospitals; [and] the reports and returns to be submitted to the Ministry by hospitals."[14]

Regulation 729 under the Public Hospitals Act describes the hospital register identifying individual patients that should be maintained by hospitals.[15] Further provisions describe the medical records necessary for each hospital patient, including medical histories and records of physical examinations and provisional diagnoses.[16] Section 38 describes in considerable detail the information that has to be included in the medical record compiled for each patient. Such extensive record keeping seems to imply an intention to use the resulting records for both research and administrative purposes. There is a similar implication to Sections 42 and 43 requiring the maintenance of certain medical records for fifty years in photographic form and establishing careful controls on the destruction of actual records or photographs; these provisions pertain to the promotion of epidemiological studies. The superintendent of a hospital is also "responsible for the safekeeping of all records relating to a patient."[17]

14 S.O., 1970, c. 378, s. 39(n)(o).

15 Regulation 729, s. 17.

16 Regulation 729, s. 37.

17 Regulation 729, s. 39(2).

Section 48 of Regulation 729 determines conditions under which a person shall be allowed to remove, inspect or receive information from a medical record. A member of the medical staff may have access for "scientific research that has been approved by the medical-staff advisory committee."[18] A Board may also permit access to medical records to "the Director of the Research and Planning Branch of the Department or his authorized representative approved by the Commission or an officer or employee of the Commission who is designated by the Chairman." Such persons may inspect and receive copies of a medical record and be given copies therefrom.[19] The regulation specifies in connection with the last provision that:

> "Subsection 6. Any information received under clause f of subsection 5 shall not be used or disclosed to any person for any purpose other than the purposes of compiling statistics and carrying out medical and epidemiological research for or approved by the Department and the Commission."

Regulation 729 is currently undergoing revision by an MOH working group, but the end results of the process are not yet certain. Changes to regulations under the Public Hospitals Act can be made by the Minister of Health with the approval of the Lieutenant Governor in Council. With respect to those aspects of Regulation 729 discussed above, the most substantial proposed

18 Regulation 729, s. 48(5)(d)(II).

19 Regulation 729, s. 48(5)(f).

modifications are in connection with Section 48. However, the draft of Section 48 is undergoing continuing revisions in the summer of 1978 and cannot yet be discussed publicly. It is to be hoped that the revised text of Section 48 will acknowledge the legitimacy of access to hospital data for research and statistical purposes.

Policy on Confidentiality and Data Dissemination

This section will discuss the policy of the Ministry of Health and in particular the Data Development and Evaluation Branch with respect to the protection of confidentiality of personal data and their dissemination for research and statistical purposes. The general provisions of the Ministry of Health with respect to the protection of confidentiality were set forth in its initial submission to the Commission on Freedom of Information and Individual Privacy in September, 1977. These are essentially identical to the "guidelines with respect to the security of identifiable patient data" developed by the Data Development and Evaluation Branch in a document dated March 23, 1978. Measures taken to ensure confidentiality in this branch are presented in Table VI. The oath of secrecy required of all civil servants is imposed by the Public Service Act and subjects

TABLE VI: MEASURES TO ENSURE CONFIDENTIALITY TAKEN BY THE DATA DEVELOPMENT AND EVALUATION BRANCH, MINISTRY OF HEALTH

In general,

1. All Civil Servants employed by the Ministry have taken the Oath of Secrecy.

2. Information pertaining to insured persons on any insured services rendered may be published by the Ministry of Health in statistical form provided the individual names and identities of persons who receive insured services are not thereby revealed.

3. Civil Servants having need to acquire or know information about an individual as a result of a contact with the health care system are required to keep that information secret unless the individual (patient) consents in writing to its release.

4. Any Civil Servant visiting a hospital to review hospital patient records on behalf of the Minister of Health has to be designated by the Minister as an "inspector" under the provisions of the Public Hospital Act. Hospitals are informed prior to any such visit.

5. It is division policy that

   a) No data is released on an individual patient to anyone else except the facility where the person receives treatment unless informed consent is given by the patient.

   b) No statistical data will be provided if there is a chance that an individual can be identified by any indirect way, i.e. residual disclosure.

   c) All other kinds of statistical information are common property and are to help facilities and government to provide better service. They are available for any useful purpose and are basic material for publication.

Source: Data Development and Evaluation Branch,
Information Services Division,
Ministry of Health, March 23, 1978.

employees to a variety of disciplinary measures for breach of secrecy.[20] The general rules on confidentiality are that no identifiable data on a particular person are released without the explicit consent of the individual. Secondly, no individual data are released in anonymized form where there is any risk of identifiability. These are in fact standard rules for releasing personal data for research and statistical purposes. The last provision in Table VI simply indicates that aggregated data can be made available to anyone, since there is no risk of identifiability.

The only difference between the guidelines of DDEB and the general rules of MOH as submitted to the Commission is a provision in the latter that "the stated policy of the Ministry on confidentiality is that data related to private persons are the property of the individual and are treated that way." This policy may have to be clarified, since the Public Hospitals Act specifically states that medical records compiled in a hospital for a patient or an out-patient are the property of the hospital.[21] This seemingly slight difference may in fact reveal an internal division within MOH concerning ownership of records

20 Public Service Act, R.S.O., 1970, c. 386, s. 10(1).

21 S.O., 1970, c. 378, s. 11.

and information derived from personal health-care records. Granting individuals some shared property rights in their own health-care data would promote concern for confidentiality on the part of custodians.

With respect to general rules on the dissemination of information, the DDEB guidelines of March 23, 1978 specify that data can be disseminated in aggregated form, as special tabulations, and in response to ad hoc requests. The publications of the branch contain only aggregate patient data. "All tables for general publication are subject to review by the professional responsible for the project and his superior. The staff members ensure that 1) no information identifying and relating to an individual is given and 2) no small cell count is published by which the individual could be identified."[22] DDEB receives standard tabulations on an ongoing basis from such information systems as medical claims and hospital admissions and discharges. Data can be abstracted from these standard tabulations to answer specific queries. On occasion the staff of DDEB themselves produce tabulations from either the main files or derived work files, unless the job complexity of the exercise

22 Data Development and Evaluation Branch, "Guidelines With Respect to the Security of Identifiable Patient Data," (Ministry of Health, March 23, 1978), p. 3. Hereafter cited as Guidelines.

requires the assistance of the staff of the Systems Management and Co-ordination Branch of the same Division: "Some of the files are also accessible to DDEB staff directly subject to the access rules which apply to SMAC personnel."[23] The access rules for DDEB determine that no file can be accessed by personnel without the written permission of the director of the branch.

Ad hoc requests for access to various types of health systems data must be submitted in writing to the director of the Data Development and Evaluation Branch. Some of these requests are submitted to a Data Dissemination Committee as discussed below. Responses to ad hoc requests can take the form of tables of aggregated data or anonymous records of individual patients. Tabulations are created from the standard existing tabulations or "from aggregating file records. The aggregation of file records may involve independent files or linked files. Any request which may breach confidentiality is referred to the Data Dissemination Committee."[24] Record abstracts of individual patients (i.e. anonymized microdata) will only be provided in response to ad hoc requests if legislation permits or patient and physician consent exists. In other cases, additional conditions or restrictions are placed on the information included in the abstract of individual data, such

23 Guidelines, p. 4.

24 Guidelines, p. 4.

as not including unique identifiers. In general, ad hoc requests can be filled by use of the medical claims file, the hospital admission and discharge system, and hospital medical records. DDEB has a procedure for requisitioning such records.

For approximately three years DDEB has had an internal Data Dissemination Committee in order to facilitate data dissemination. The director of DDEB sends only the difficult or sensitive requests for access to data to the committee. In particular, it reviews "any request which involves potential or real breach of confidentiality and/or high production cost."[25] The membership of the committee is at the working level, including the four DDEB section chiefs, the systems co-ordinator from SMAC Branch involved with information retrieval, a chairperson, and an economist. The committee meets as a whole at least once a month and more frequently if required. The usual sequence procedure for a researcher seeking access to MOH data is a telephone call to DDEB, a personal interview at the branch, and, possibly, an actual meeting with the Data Dissemination Committee, which can set the conditions for a particular dissemination.

25 Guidelines, p. 4.

The Data Dissemination Committee may recommend fulfilling a request, restructuring it, or placing conditions on the release of data. It handles approximately 25 to 30 requests per year. For the nine months ending in June, 1978 the committee processed 17 requests. The committee granted permission to release the data in 53 percent of the cases, required further clarification or studies in 29 percent of the cases, and rejected 3 requests (18 percent). Minutes are kept of the monthly meetings of the committee. The main subjects recently reviewed at meetings pertained to confidentiality and the recovery of data retrieval costs.[26] The committee designed and used forms or letters to resolve any problems of confidentiality relating to OHIP individual data. It has also analyzed the cost of data retrieval distribution by operational areas and by range of payments. Rules for recovering costs are being finalized.

The three requests for access that were turned down in the period ending June, 1978, are instructive of the problems faced by the Data Dissemination Committee. In the first instance the Ontario Economic Council for a study directed by Pran Manga appeared to want OHIP numbers and addresses for particular

26 Much of the information on the work of the committee is derived from a report on its activities furnished by the director of the Data Development and Evaluation Branch.

individuals. A meeting between the committee and the researchers determined that these particular pieces of information were not indeed required, and that anonymized microdata would be satisfactory. In the second episode a graduate student from the University of Toronto working on a doctoral dissertation sought a duplicate of some tapes furnished to the Ontario Economic Council in connection with the Manga studies in order to elaborate on and re-use the data. When the student could not afford production costs of $750., the committee suggested that the tapes of anonymized microdata be borrowed from the Council. This transfer did not occur because the merged tapes were not then in a readily-available form. The third refusal involved researchers at the University of Toronto seeking tabulations of anesthetic procedures used for specific types of operations at specific hospitals. In this instance the MOH simply did not have the data.

In general the Data Dissemination Committee will not deny aggregate data to anyone. It will also not release identifiable microdata without the consent of the respondent. Table VII presents a consent form used by patients to authorize a research team from the University of Toronto to examine their OHIP records for a study of the relationship between fitness and worker productivity.[27]

27 See below pp. 89-91.

TABLE VII: FITNESS HEALTH COST RELATIONSHIPS
CONSENT TO ACCESS OHIP RECORDS

As part of the study, the research team wish to find out the cost of providing all the OHIP services you have received since April, 1972. To do this the team wish to extract the record of claims made by your physicians for treatment they have given you. (This record states the category under which each claim was made, eg. X-Ray, check-up, etc.). The research team need your consent to extract this record. The information will be used in the strictest confidence and only in a way which makes it impossible for anyone outside the research team to identify your record.

A. If you consent to the research team examining your record, please sign below:
I consent to the research team examining my record of OHIP claims from ..... to .....

Signed: Date:

Please now complete the back of this form.

B. If you do not consent to the research team examining your record, please sign below:
I do not consent to the research team examining my record of OHIP claims.

Signed: Date:

Please do not complete the back of this form.

NAME (Please Print) OHIP #

DATE OF BIRTH _____ AGE _____
Day Month Year years

If you are female and were married after April 1972, please give your name before marriage. _____

Please give your type of OHIP coverage from April 1972 onwards - as best you can remember.

FROM	TO	SELF ONLY	TYPE OF COVERAGE INCLUDED WITH FAMILY

Source: Data Development and Evaluation Branch, Information Systems Division, Ministry of Health.

Some problems do exist in terms of obtaining access to Ministry of Health data for research and statistical purposes, since the pertinent legislation does not explicitly or adequately authorize access by outside researchers. Despite the provisions in Regulation 729, for example, some critics question the legality of access by outsiders to MOH data for research. They claim that the legal authority for such access is inadequate. It is asserted that research is not legally permitted under the Health Insurance Act and that some of the things done by researchers are illegal under the terms of the Public Hospitals Act. For example, a McMaster University project had the consent of a hospital to study the quality of primary care in the hospital, but researchers are not permitted to examine hospital records under the Public Hospitals Act, nor can medical doctors on the staff examine them at will. The Ministry of Health finally had to permit this particular research project under special permission from the Minister.

A second problem for researchers is learning what data are available for research purposes from MOH. The Catalogue of Statistical Files in the Ontario Government 1977, which lists a large number of MOH data files but is not an MOH product, is for internal government use only. The Catalogue does list the objectives, content, accessibility, and size of a particular file. Although the volume is described as a catalogue of statistical

files, some of the objectives of the data files are clearly administrative in nature, such as "to determine the medical necessity of a patient remaining in hospital."[28] In some instances the Catalogue indicates that one objective of a file is "to supply information for research studies."[29] An outsider seeking access to MOH data through the Catalogue might also be mystified by the varying characterizations of the "accessibility" of personal data. Of 17 data files in the 1977 Catalogue that appear to contain personal data, 11 are described as strictly confidential, 3 as confidential, 1 as confidential to some extent, 1 as "not stated," and 1 as not confidential. A perinatal mortality survey contains data related to circumstances of birth (and death, if applicable) for each child in each of ten hospitals for the years 1960 and 1961. The purpose of the file is to study factors related to perinatal death in Ontario. There are 51,490 births in the cohort. The accessibility is described as "confidential to some extent, e.g. details of maternal deaths and specific hospital." It is unclear whether the specific births are identifiable as such. The mode of storage is described in peculiar terms as well: "Ministry

28 Catalogue, p. HL15.

29 Catalogue, pp. HL19, HL20.

magnetic tape at University of Western Ontario, Johns Hopkins University, and Data Development and Evaluation Branch."[30]

The characterization of accessibility "not stated" is used in the Catalogue in connection with a file of mental health data held by the SMAC Branch of Information Systems Division. The contents appear to be designed for statistical reports, but contain a yearly file of admissions and separations to and from psychiatric hospitals and units and residential units. There is also an annual census file for each type of facility. The identifiers appear to be birthdate and casebook number and sex.[31]

The characterization in the 1977 Catalogue of "not confidential" applies to a file entitled "Communicable Diseases." It is held by the Community Health Protection Branch of the Community Health Division. The stated objective of the file is to study the incidence and prevalence of any communicable disease. It contains "records of all cases of some twenty communicable diseases including typhoid, rabies, whooping cough, and leprosy."[32] There are 100,000 items in this file identified by the names of

30 Catalogue, p. HL22.

31 Catalogue, p. HL28.

32 Catalogue, p. HL4.

cases of each disease. It is unclear from the Catalogue whether or not specific individuals are identifiable.

A third problem for outside researchers in dealing with DDEB concerns their allegedly low order of priority in the hierarchy of branch and government activities. This is coupled with an alleged reluctance to let information leave the branch and the low budgets of outside researchers in the face of the current trend of the branch to recover costs. It seems likely that the actual situation primarily reflects the predictable problems encountered by any outside researchers attempting to obtain access to data held by a busy administrative agency. DDEB procedure does encourage researchers to visit the branch and see what data are available before various forms of releases are requested.

The final concern in using MOH personal data for research and statistical purposes is of a general nature. The various information systems in the ministry were set up to pay physicians and hospitals and not to track patients or to link data on individuals. The lack of a Universal Personal Identifier (UPI) or Personal Identification Number in Ontario severely restricts the utility of OHIP data for research and statistical purposes. Although some record linkages are accomplished for internal and external uses, and although some

work is being done on overcoming internal obstacles to record linkages, the lack of a UPI severely limits this practice. For example, in a feasibility study for the Health Care Unit of the University of Toronto, it was too difficult for a research project to link accident reports with ambulance, hospital, and mortality data in order to evaluate the adequacy of services. Record linkages can only be undertaken at great difficulty and expense by means of the Soundex system. As will be discussed further below, the absence of a UPI severely limits the use of MOH data for epidemiological studies. Thus it is generally recognized both within and outside the ministry that the introduction of a UPI is essential. In its submission to the Commission on Freedom of Information and Individual Privacy in September, 1977, the Ministry of Health stated its intention to introduce a health care number, which would be provided to every citizen of the province: "The decision has been taken by government to base this Health Care Number upon the federal Social Insurance Number and we are satisfied that this is the most appropriate number, not only for Ontario but a selection which should assist all Canadians in some way in achieving better health care, not only from program delivery but also medical research." The Minister of Health, Dennis Timbrell, has also publicly stated that he has Cabinet support to go ahead with the unique form of plastic card that will contain a patient's Social Insurance Number.[33]

33 London Free Press, May 6, 1978, p. A3.

Uses of Ministry of Health Data

This section will illustrate how researchers have used MOH data on an ad hoc basis for health-related research. The first area of concentration will be the research activities of health economists sponsored by the Ontario Economic Council. The Council is an independent research institute that studies public policy problems in the area of natural resources, human resources, government, and provincial economic development. It is publicly-supported by the Ministry of Treasury, Economics, and Intergovernmental Affairs. The Ontario Economic Council Act permits the Council to "conduct socio-economic studies in any area considered by the Council to be of concern; [and to] cause to be published such studies and reports as are prepared by or for the Council."[34]

An initial example of work in the health field sponsored by the OEC is a study undertaken by Dr. Morris Barer, a health economist on the staff of the Council. He is developing various measures of "case-mix complexity" for hospitals and measuring the impact of case-mix variations on unit cost variation. The complexity measures are based on a provincial distribution of patient discharges from hospitals. By means of a written

34 Ontario Economic Council Act, S.O., 1970, c. 309, s. 5(a)(b).

request to the Data Dissemination Committee of the Data Development and Evaluation Branch, Barer in 1977 obtained abstracted portions of individual hospital discharge records in the form of anonymized microdata.

A second example of health research in progress sponsored by the Ontario Economic Council is a study of direct-billing of patients by physicians undertaken by Professors Alan Wolfson and Carolyn Tuohy of the University of Toronto. The study compares the behavior of Ontario physicians, who have chosen to remain out of the OHIP insurance scheme with that of their peers within the OHIP system, "in terms of levels of services provided, the mix of services, their use of other health resources, cost per patient, encounters per patient, characteristics of physicians in terms of sex, date or place of graduation, and practice-type, and so on."[35] Much of the relevant data were obtained through a survey of Ontario physicians conducted by the researchers. But in order to study the "political economy" of direct-billing physicians, Wolfson and his associates had to obtain some data from the Ministry of Health. Because of concern for the confidentiality of responses to sensitive questions in their survey, the researchers did not want to give the survey tape directly to the MOH. Instead they

35 Ontario Economic Council, Annual Report 1975-76, (Toronto, 1976), p. 14.

furnished the survey data in question to the physician caring for particular patients; the physician then obtained profiles on his practice from the MOH, removed identifying particulars about individual patients, and gave the merged data on individuals to the research team in anonymous form.

The Council has just published a study by Pran Manga of the University of Ottawa of the distributional impact of health care financing in Ontario. This study investigates the distribution of the benefits of Ontario's medical insurance plan by income class.[36] It is essentially an effort to determine the extent to which different economic segments in the population use the health care system.

The Council has now asked Professor Manga and Dr. Barer to extend the initial Manga study to look at hospital use by income class and to develop a model of the use of hospital services in Ontario. The research will address two key questions: "First, who utilizes and benefits from the hospital insurance program; second, what factors appear to determine or

36 See P. Manga, The Income Distribution Effect of Medical Insurance in Ontario, Occasional Paper 6, (Ontario Economic Council, Toronto, 1978).

explain variance in hospital utilization across Ontario families and individuals."[37] Since acquiring answers to such questions is obviously in the public interest, one of the fascinating aspects of the Manga and Barer research projects has been their difficulties in constructing a satisfactory data base from which to carry out their analyses. Given the expenditure of four billion dollars annually on health care in Ontario, it is surprising that it is so difficult to obtain data to study the system. This is in part because MOH only collects health-related data. Manga made the following report in connection with his first study:

> "The major difficulty faced in the estimation of benefit incidence is the absence of the required data base. Thus a survey has been undertaken to develop a microdata base including the important socio-demographic and income information on a sample of OHIP families and some pertinent health-related data. The survey data are then merged with the utilization records of the Ministry of Health to generate the complete base for the estimation of the benefit incidence exercise."[38]

For the follow-up study the researchers require hospital discharge data for families participating in that survey. Such data are in the custody of the Ministry of Health through the

37 M. Barer and P. Manga, "A Benefit Incidence and Utilization Analysis of the Hospital Insurance Program in Ontario: Methodology and Some Preliminary Results," A Paper presented to the Canadian Economics Association, London, Ontario, May 30, 1978, pp. 1-2.

38 Ontario Economic Council, Annual Report, 1975-76, p. 14.

OHIP system. But the MOH medical and hospital files are not designed for access to and linkage of individual patient data. They are not stored for access by researchers but to facilitate administrative functions. The data thus have to be transformed and linked in some manner for the researchers.

Thus even before confronting such fundamental issues as the legal and ethical aspects of confidentiality or the sanctity of the physician/patient relationship, researchers have significant problems in obtaining access to data. If they wish to obtain a sample of families belonging to the OHIP system, no specific sampling frame exists listing the OHIP families, but one can be developed for a specific project. Secondly, the hospital discharge data at MOH contain no information on such important variables as family size, family income, and education/occupation of the head of the family or spouse. Sample surveys have to be undertaken to develop such information on a voluntary basis from the individuals themselves. Thirdly, the accessibility of records by use of the OHIP number produces obvious problems of confidentiality for the researchers and for the ministry. The latter will not release individual data with OHIP numbers attached. In addition, the OHIP number is for an entire family, so that the researchers have to use age and sex data on members of a family to establish a file of individuals. In addition, before April, 1972, many families and individuals did not have

the same number for both hospital and medical care services, which obviously hinders longitudinal studies. Without a Unique Personal Identifier the researchers have no idea of the size of particular families. The absence of such a numbering scheme also makes it impossible to link data on use of the medical system and income class.

In the Manga study an interview survey was used to develop a sample of OHIP families in order to study their use of medical and hospital services during a twelve-month period in 1974-75. The interviews conducted by York University's Survey Research Centre provided OHIP numbers and a variety of socio-economic information on families. The researchers gave the survey data tape to the Data Development and Evaluation Branch in identifiable form for purposes of linkages with OHIP data. The linkages were done by computer for the Ministry of Health at the expense of OEC. The researchers obtained in return anonymized microdata without identifiers such as names, addresses, or OHIP numbers.

Although such an arrangement has proved satisfactory for the Manga and Barer studies, its somewhat cumbersome and time-consuming characteristics appear obvious, even for researchers with a direct association with a government research institute. Yet it is also clear that access to microdata on individuals for

various types of research, including health research, is becoming a growing need.[39] Researchers need to obtain access to statistical tabulations or anonymized microdata based on linked records. Thus it is evident that the OHIP information systems should be rationalized in order to facilitate multiple uses of data, including the needs of the research community. If, as appears evident, research in health economics is in the public interest, health economists should not be able to complain that the time required to obtain access to data is their biggest problem. At the same time researchers must recognize that certain time delays are almost inevitable and that services rendered by the MOH cannot always be free. DDEB can fulfill a straight, simple request for a tabulation within a week, but complex tabulations involving refinements can take two to four months. The research community has to plan their activities with such realities in mind and approach the custodians of data in a sensitive and accommodating manner. It should also be understood that some of the delays are the result of necessary concern for confidentiality in preparing data for release.

Another example of a major research project requiring access to Ministry of Health data is the study of "employee fitness and worker productivity" currently being directed by Professor Roy J. Shephard, Professor of Applied Physiology, and Michael H. Cox, a

39 See David H. Flaherty, Privacy and Government Data Banks. An International Perspective (New York, forthcoming, 1978), passim.

research associate in the Department of Preventive Medicine and Biostatistics, at the University of Toronto. Health and Welfare Canada (Fitness and Amateur Sport Section) is funding the study of the impact of fitness on productivity. The objectives of the study are as follows: 1) to define current fitness levels in a selected subpopulation of "healthy" adult office workers; 2) to determine by controlled trial whether current fitness levels of the adult worker relate to job productivity, and whether productivity can be increased by well-regulated employee fitness programs; 3) to determine the relationship between fitness levels and health costs, and to test whether health costs can be reduced by participation in an employee fitness program; and 4) to examine the possible contribution of an employee fitness program to job satisfaction and general well-being in the adult population. The researchers assert that in view of the economic importance of the topic, it is surprising that there has been almost no research on this series of issues. The present study compares the employees of two large Canadian life insurance companies in the Toronto area, which have a significant interest in preventive medicine. The methodology involves measures of productivity, including health experience, psychological measures, and physiological measurements. Approximately 1,000 individuals from the two companies are involved. With promises

of confidentiality to the individuals involved, the researchers conducted physiological stress tests and life style analyses. The respondents have been very cooperative.

In order to measure health experience, the researchers had to obtain some data on employees from the Ministry of Health. They sought individual data on health costs, diagnosis codes, and visits to doctors. Almost all respondents voluntarily furnished their OHIP numbers to the researchers. When the controversy over RCMP access to OHIP data arose in the fall of 1977, the ministry was reluctant to release the data. The researchers initially sought grouped OHIP data on individual employees in terms of four categories of OHIP information: total annual OHIP payment; hospital length of stay; hospital discharge diagnosis; and fee service code. At a cost of approximately $13,000, the Information Systems Division finally agreed to release anonymized microdata for approximately 650 persons involved in the study. Such microdata will be in fact even more useful to the researchers than the grouped data that they originally sought. The researchers did agree to protect confidentiality at the time of the release of the data.

A good example of the type of ad hoc need for access to individual records at the Ministry of Health arises from some of the current research of Professor E. Aileen Clarke of the Department of Preventive Medicine and Biostatistics of the University of Toronto. In the first example the Ontario government was recently concerned that a new agent in the environment was causing an increase in acute lymphatic leukemia in an area of southern Ontario where the substance was used. The Ontario government asked Dr. Clarke to study whether the incidence of this type of disease had in fact increased in the particular area. Because of her formal position as head of the Division of Epidemiology and Statistics for the Ontario Cancer Treatment and Research Foundation (OCTRF), Dr. Clarke was able to use mortality files on cancer through 1977, as well as other sources of information.[40] She was then able to obtain permission from the Ministry of Health to use the recent hospital files to obtain current data. She reported the rsults of her research to the government on May 15, 1978, which showed that there had not been an increase in leukemia in this particular area in recent years.

40 See Table XIV.

In a second more intensive study, Dr. Clarke and Professor T.W. Anderson are studying carcinoma of the cervix. With permission from OCTRF and the hospital, she was able to obtain the names of a group of patients of the Princess Margaret Hospital in Toronto who were suffering from this disease. With the permission of individual physicians, she conducted interviews with 221 patients, asking them in particular about the number of Pap tests that they had had during the last ten years. She was able to obtain the same information from a control group of women from the same neighborhood who did not have the disease. She found a significant difference in results between the patients suffering from the disease and the control group, but was unable to interview all of the members of the patient sample suffering from the disease because of language difficulties or the illness of the patient. Her solution was to furnish the Ministry of Health with the OHIP numbers of those patients that she had been unable to interview; she asked for statistical information on those that had had no Pap smears versus those that had one or more. The Ministry of Health produced statistical data in grouped form that did not identify particular patients. This would appear to have been a legitimate research and statistical use of OHIP data, despite the fact that there has been some informal criticism of Dr. Clarke's work in this particular instance because she did not have specific informed consent from these individual patients to obtain OHIP data. For those

actually interviewed, "permission was given by 95% of the cases and 85% of the controls for the investigators to have access to their medical records."[41]

Professor David Hewitt of the Department of Preventive Medicine and Biostatistics of the University of Toronto has had the frustrating experience of discovering that research conducted in the 1960s and early 1970s was no longer possible. Table VIII lists a series of seven articles by Professor Hewitt and various associates. They cover such topics as the desirability of incidental appendectomy, the prevalence of hemophilia, the practice of early discharge of maternity patients, the occurence and treatment of gall bladder disease in Ontario, cancer morbidity, and breast cancer. Each of these seven articles involved direct access to identifiable hospital discharge records held by the Ontario Hospital Services Commission, which was the forerunner of OHIP. Two of the articles on cancer also involved access to identifiable data held by the Ontario Cancer Treatment and Research Foundation, which swore Hewitt to secrecy under the Cancer Act. Hewitt had free access to the hospital discharge records in the custody of the Hospital Services Commission,

41 Ontario Cancer Treatment and Research Foundation, Cancer in Ontario, 1977 (Toronto, 1977), p. 199.

TABLE VIII: PUBLICATIONS OF PROFESSOR HEWITT OF THE UNIVERSITY OF TORONTO INVOLVING ACCESS TO HOSPITAL DISCHARGE RECORDS FOR ONTARIO

1. David Hewitt, M.A. Jean Milner, and W.H. LeRiche, "Incidental Appendectomy: A Statistical Appraisal," Canadian Medical Association Journal, Vol. 100, June 21, 1969, pp. 1075-1081.
2. Hewitt and Jean Milner, "Prevalence of Hemophilia in Ontario, 1966," Canadian Medical Association Journal, Vol. 102, January 31, 1970, pp. 174-77.
3. Hewitt, Anne Pettypiece, and Jean Milner, "Early Discharge of Maternity Patients," Canadian Hospital, July, 1970, pp. 44-48.
4. J. Milner and D. Hewitt, "The Occurence and Treatment of Gallbladder Disease in Ontario," Journal of Chronic Disease, 1972, Vol. 25, pp. 73-83.
5. David Cook, E.N. Mackay, and D. Hewitt, "Cancer Morbidity in National Origin Subgroups of the Ontario Population," Canadian Journal of Public Health, Vol. 63, March/April, 1972, pp. 120-24.
6. D.C. Cook, O. Dent, and D. Hewitt, "Breast Cancer Following Multiple Chest Fluoroscopy: The Ontario Experience, CMA Journal, Vol. 111, September 7, 1974, pp. 406-410.
7. D. Hewitt and J. Milner, "Inequality of the Services Received by Individuals: A Suggested Index and an Illustrative Application," Medical Care, Vol. 13, No. 11, November, 1975, pp. 928-933.

because he was attached at that time to the Metropolitan Toronto Hospital Planning Council, a government body located at Commission headquarters. He was at the same time a faculty member of the University of Toronto. These seven articles were in fact "a series of studies conducted under the auspices of the Metropolitan Toronto Hospital Planning Council and made possible by access to records of the Ontario Hospital Services Commission."[42]

In the first of this series of articles, Hewitt discussed the possibility that loss of appendix may contribute to an increased expectation of malignant disease, especially cancer of the colon.[43] The article also pointed out that existing records of the Hospital Services Commission were ideal to study this "scientific question of some importance." "We therefore hope to see a file maintained with identifying information on subjects undergoing incidental appendectomy, and on comparison groups not so treated, in a form that would permit linkage to future records of diagnosis, treatment and death from malignant disease."[44] When Hewitt attempted to obtain access to the relevant records,

42 David Hewitt, M.A. Jean Milner, and W.H. LeRiche, "Incidental Appendectomy: A Statistical Appraisal," Canadian Medical Association Journal, C (June 21, 1969), 1081.

43 Ibid., p. 1080.

44 Ibid., p. 1080.

he encountered Regulation 729 under the Public Hospitals Act, which required him to be recognized as a member of the research staff of the Ministry of Health and to obtain the approval of the ministry itself. It took six months for him to become an "authorized representative" of the research branch of MOH; after another six months he had been unable to have the ministry itself approve his project. Since the exercise of attempting to obtain access had occupied approximately one year and his research funds were drying up, Hewitt determined to use the existing funds for another research purpose. Hewitt has not tried recently to obtain access to identifiable hospital discharge records. Other researchers have obtained access under the existing regulations of the Public Hospitals Act by writing to MOH and being designated by the Minister as a representative. This general solution is currently available to outside researchers under the regulations of the Public Hospitals Act.[45]

Hewitt points out that each of the articles listed in Table VIII involved the use of data containing some identifiable particulars of persons ascertained through hospitals or medical treatment records or as covered by medical services insurance: "In none of these instances did the research involve any reference to the named subjects for their assent to inclusion in the study, and

45 Regulation 729, s. 48(5)(f).

in no instance were these people asked to participate in the study in any additional way." But Hewitt now finds that this type of research seems impossible. He is also of the opinion that a methodology or protocol must be developed whereby researchers in the health field can approach respondents directly to discover whether or not they are willing to participate in follow-up studies. Hewitt found from his earlier experience that the overwhelming majority of patients and their relatives are quite agreeable to disclosure and hospitable to researchers who want to ask them supplementary questions: "But because we have a system set up to protect the anonymity of such people, there is no legal way for an investigator to ascertain whether a particular patient would agree to disclosure or not." In his view a method is needed by which the wishes of the individual patient or his physician can be ascertained, so that we do not have to live with the invariable assumption (wrong more than 99 percent of the time) that these persons would not wish to assist research by permitting access to their records.[46]

46 Hewitt has kindly made available a written statement and copies of some of his correspondence on this issue.

Recommendations

The recommendations in this section are based on both the recent experience of this writer in examining the Ministry of Health and the needs of researchers in Ontario, and previous experience in studying similar issues at the national and international level. The product of the latter research will not be reproduced in detail in the current effort. It should be understood, however, that between 1974 and 1978 this writer undertook an extensive study of the collection and dissemination of personal data by national government agencies in five countries, including Canada, the United States, Great Britain, West Germany, and Sweden.[47] A similar methodology has been followed in writing about health-related agencies in Ontario in the current presentation. The volume treats both agencies created to collect statistical data and those that use administrative data for statistical purposes. The summary recommendations for government agencies derived from this research are presented in summary form in Table IX.

The five case studies were also used as background reading for the Bellagio Conference on Privacy, Confidentiality, and the Use of Government Microdata for Research and Statistical Purposes,

47 Flaherty, Privacy and Government Data Banks.

TABLE IX: SUMMARY RECOMMENDATIONS FOR GOVERNMENT AGENCIES

1. Establish policy, rules and regulations concerning data protection, confidentiality, and the dissemination of microdata.

2. Devote considerable efforts to explaining to the general public the procedures in force for the protection of the confidentiality of individual information collected and used for research and statistical purposes. Publicize and explain the intended uses of individual data for research and statistical purposes.

3. Encourage and assist researchers by disseminating as much data as possible within the confines of confidentiality.

4. Formulate precise rules and regulations for granting access to data.

5. Establish advisory committees to assist in the formulation of policies on confidentiality and the dissemination of data.

6. Prepare public use samples from censuses and social surveys.

7. Encourage record linkage whenever it is determined to be in the public interest.

8. Follow a careful protocol in preparing a sample for research. A department should seek informed consent from a particular respondent before permitting a researcher to contact members of a sample.

9. Require written undertakings for the protection of confidentiality from users of sensitive data. These should specify in explicit detail the research which is to be undertaken and the uses to which the data can be put.

10. Encourage professional organizations to formulate codes of ethics for their members with respect to the use of individual data for research.

Source: David H. Flaherty, Privacy and Government Data Banks. An International Perspective (Science Associates/ International, Inc., New York, forthcoming, 1978), chapter 21.

held in Italy under the auspices of the Rockefeller Foundation in August, 1977. Eighteen representatives of national government agencies and users of individual data for scientific purposes met for a five-day period. They developed a set of eighteen principles to govern data collection, protection, and dissemination for such purposes. The Final Report of this conference is reproduced as Appendix One. Again it should be repeated that the findings of Privacy and Government Data Banks and the Bellagio Conference should, whenever possible, be applied to the health research field in Ontario. The following discussion will emphasize more specific recommendations for the Ministry of Health in Ontario.

1. The Ministry of Health should continue to promote and encourage the dissemination of health-related data for research and statistical purposes. The Data Development and Evaluation Branch has already made significant contributions to promoting scientific uses of MOH data. The series of separate information systems that currently exist should continue to be integrated to the extent necessary to promote scientific uses of MOH data. The public has the right to expect greater use of and access to data in order to learn more about the benefits of the substantial current spending on health care in Ontario. Problems in obtaining access to data for legitimate purposes are part of the

reason why not enough is currently known about such questions. The OHIP record system must be rationalized in order to promote multiple uses, including research and statistical access to administrative data bases.[48] Such activities are very much in the public interest. In order to promote research uses, the Information Systems Division in particular should make available to outsiders a basic guide or catalogue to the main data holdings of the division. The end result of such continuing actions on the part of the Ministry of Health should be strengthened support for health research. Any current legal provisions that hinder access to data for approved health research should be changed to recognize the legitimacy of research uses of data. In order to promote access to data, it might also be desirable for the Data Development and Evaluation Branch to have several outsiders on its Data Dissemination Committee in order to ensure that refusals of access occur for legitimate scientific and/or ethical reasons.

2. The Ministry of Health should establish (or at least formally recognize) a separate health statistics centre within the Information Systems Division. Thi would permit implementation of the principle of a strict functional separation

48 See, generally, A.J. Culyer, Measuring Health: Lessons for Ontario, Ontario Economic Council Research Studies (University of Toronto Press, Toronto, 1978).

between administrative/regulatory and research/statistical uses of personal data. In order to promote concern for confidentiality and access by outsiders to health-related data for scientific purposes, it is essential that there exist an internal division between the use of personal data for administrative/regulatory and research/statistical purposes. The essential distinction is that an administrative use of personal data may have a direct impact on a particular person, whereas research and statistical uses of data result in information about groups of individuals. A legitimate use of personal data for research and statistical purposes should never result in action directly against a particular person, unless such a use is built into the research protocol, as in the case of various forms of treatment research. The Information Systems Division currently seems to blend administrative and statistical uses of data in a somewhat unsatisfactory manner. For example, the catalogue of government data files lists a blend of administrative and statistical data bases in the Information Systems Division.[49] These listings at least suggest that the two branches of the division engage in both administrative activities, such as monitoring the practices of physicians, and purely evaluation research and statistical activities. If the division is only supposed to do statistical and evaluation research, as opposed to actual monitoring of

49 Ministry of Treasury, Economics, and Intergovernmental Affairs, Catalogue of Statistical Files in the Ontario Government, 1977 (Toronto, 1977).

particular persons, then a stricter implementation of the principle of functional separation is needed. Having a separate health statistics centre would also make it possible to avoid such thorny problems as defining a researcher; in this instance a researcher would be defined as anyone seeking access to data from the health statistics centre and willing to abide by its rules on confidentiality.

Advocating the creation of a single health statistics agency within the Ministry of Health was the basic thrust of a 1975 report by the Ontario Council of Health on Health Information and Statistics.[50] The Ontario Council of Health is the senior advisory body to the Minister of Health on health matters. This particular report was prepared by the Committee on Health Information and Statistics and submitted to the Council for approval on November 26, 1974. The report found the current health statistics system in Ontario inadequate; "Despite the variety of sources and the quantity of data available, Ontario's health statistics system embodies some major weaknesses that limit its usefulness to planners, managers, and researchers in the health care system. Many of these shortcomings were identified in 1969 by the Committee on Health Statistics."[51] Although the Council

50 Ontario Council of Health, Report on Health Information and Statistics, 1975 (Toronto, 1975).

51 Ibid., p. 6; see also Ontario Council of Health, Report on Health Statistics. Part I. Annex "G". January, 1969 (Toronto, 1969); and Part 2. Implementation of a Health Statistics System, Supplement No. 2, 1970 (Toronto, 1970).

praised the Ministry of Health for creating the Information Systems Division, it still believed that a variety of deficiencies "point to the need for a more efficient, comprehensive, and co-ordinated health statistics system in Ontario."[52] The Council of Health concluded that "responsibility for the planning, co-ordination, and direction of the health statistics system in Ontario should be vested in a single health statistics agency within the Ministry of Health."[53] The report included an elaborate discussion of the proposed functions of such a health statistics centre.

In terms of establishing guiding principles for an efficient and effective health statistics system in Ontario, the report asserted the principle that "the system should be so designed as to prevent unauthorized access to information, so that the confidentiality of personal health information is respected and the privacy of the individual is protected."[54] The discussion of this principle included a statement that the public should be informed of all rules and standards of access to and disclosure of personal health information. The Council of Health also

52 Health Information and Statistics, p. 16.

53 Ibid., pp. 6, 21.

54 Ibid., p. 20.

recommended that the proposed health statistics agency "should have the authority and the responsibility to impose such controls and constraints on data access as may be required to protect the confidentiality of personal health information and to preclude invasion of the privacy of the individual."[55] It would probably be desirable for such authority to be established by either statute or regulations.

The proposed health statistics centre would have the responsibility to promote research and statistical uses of data in the scientific community. The Council proposed that the agency "should develop mechanisms for screening requests for health data, determining their urgency and priority in relation to other statistical needs, and maintaining a continuing record of unmet requests throughout the system."[56] The Council also made the useful suggestion that the Ministry of Health should 'establish' an Advisory Committee on Health Statistics, whose membership would be broadly representative of producers and users of health statistics within and outside the Ministry."[57]

55 Ibid., p. 26.

56 Ibid., pp. 23-24.

57 Ibid., p. 26.

3. One of the most important new innovations in data dissemination for research and statistical purposes in recent years has been the preparation of public use samples of personal data. This methodology was pioneered by the American Bureau of the Census, adopted by the National Center for Health Statistics in the United States, and recently followed by Statistics Canada.[58] A public use sample consists of perhaps a one percent sample from an existing data file of anonymized microdata. No particulars identifying a person uniquely are included on the computer tape. The Ministry of Health should investigate the possibility of preparing public use samples for medical claims files and hospital claims files generated through the OHIP system. The geographical identification on such tapes could be provincial in nature with further geographical identification for areas with at least 250,000 persons. The latter number is the traditional figure used in existing public use samples from censuses and social surveys. It may not be necessary for the Ministry of Health to restrict itself to a one percent sample. The federal Ministry of Revenue makes available a six percent sample of Ontario records for individual taxpayers. The model of public use samples published by the National Center for

58 Flaherty, Privacy and Government Data Banks, Parts IV and V.

Health Statistics in the United States seems particularly relevant to the work of the Ministry of Health.[59]

4. Perhaps the single most important innovation that the government could make to promote the use of health data for research and statistical purposes would be the adoption of a system of Unique Personal Identifiers for residents of Ontario. The deficiencies of the current system of assigning OHIP numbers by families has already been discussed. Since the 1960s, individuals within the Ministry of Health and outside commentators have been advocating the utility of a UPI in order to promote the use of data.[60]

The Ministry of Health has already informed the Commission on Freedom of Information and Individual Privacy of its intention to introduce a Unique Personal Identifier for the health system. The intention is to adopt the Social Insurance Number of the federal government. From the point of view of promoting health

59 U.S. Department of Health, Education and Welfare. Public Health Service. National Center for Health Statistics, Standardized Micro-Data Tape Transcripts. DHEW Pub. No. (HRA) 76-1213. (Rockville, Maryland: National Center for Health Statistics, February, 1976); and National Center for Health Statistics, Policy Statement on Release of Data for Individual Elementary Units and Special Tabulation. DHEW Pub. No. (PHS) 78-1212. (Hyattsville, Maryland: National Center for Health Statistics, May, 1978).

60 See Ontario Council of Health, Report on Health Statistics, 1969, pp. 5, 9, 10; Ontario Council of Health, Implementation of a Health Statistics System, 1970, p. 30.

research, and even taking into account the risks of administrative abuses of such a unique identifier, it is the opinion of this observer that a UPI is very much in the public interest, so long as controls on its use are built into the system and enforced.[61]

5. Having adopted a Unique Personal Identifier, the Ministry of Health should promote internal record linkages for research and statistical purposes. It should also devote particular attention to promoting epidemiological research on occupational and environmental hazards in Ontario.[62] Advocacy of medical record linkage has had a long history in Canada. Dr. Howard B. Newcombe of Atomic Energy of Canada has been a pioneer at the international level for the last twenty years in developing methods of record linkages.[63] An Ad Hoc Committee of the Medical Research Council of Canada produced a major report in 1968 entitled Health Research Uses of Record Linkage in Canada.[64]

61 For an extensive discussion of the risks of abuses of personal data in a variety of national contexts, see Flaherty, Privacy and Government Data Banks, passim.

62 For a comparable recommendation in connection with Saskatchewan health-care data, see C.A.R. Dennis, et. al., Medicare Data: Its Use in Defining the Effects of the Environment on Health (Prairie Institute of Environmental Health, Regina, October, 1975).

63 For a listing of the many publications by Newcombe, see David H. Flaherty, E.H. Hanis, and S. Paula Mitchell, compilers, Privacy and Access to Government Data For Research, An International Bibliography (Science Associates/ International, Inc., New York, 1978), pp. 70-72.

64 Medical Research Council of Canada, Health Research Uses of Record Linkage in Canada, Report No. 3 (Ottawa, October, 1968).

The Ontario Council of Health and its Committee on Health Statistics supported the idea of record linkages in its 1969 and 1970 reports on health statistics.[65]

The Ontario Council of Health's 1975 report on Health Information and Statistics strongly supported a system of record linkages in the health field and the promotion of epidemiological research.[66] Its "proposed health statistics agency should have the potential and the capacity for linkage of health records and reports."[67] The need for record linkage was described in detail.[68] The Council argued that one of the guiding principles for an efficient and effective health statistics system in Ontario should be that "the system should provide a potential for the linkage of information from multiple sources."[69] The goal will be to link information from multiple sources that exist in Ontario so as to make maximum use of available data for scientific research. The Council regarded the promotion of

65 Ontario Council of Health, Report on Health Statistics, 1969, pp. 5, 9, 10; Implementation of a Health Statistics System 1970, p. 30.

66 For comparable conclusions in another excellent report, see Ontario Council of Health, Health Care Delivery Systems. The Role of Computers in the Health Field. Supplement No. 9 (Toronto, 1970).

67 Ontario Council of Health, Health Information and Statistics, pp. 25-26.

68 Ibid., p. 24.

69 Ibid., p. 20.

improved linkage of health data as a priority area for early study and implementation by its proposed health statistics agency.[70] Another guiding principle should be that "the system should facilitate the conduct of epidemiological research and other types of health research."[71] In every instance the Council elaborated on the rationale for such recommendation in the pages of its report.

The Ontario Council of Health through its Committee on Health Research Development returned to similar themes in its 1977 report on Health Research Priorities for Ontario, which also emphasizes the importance of record linkages and epidemiological research. Consultant groups in epidemiology and biostatistics informed the Council through its Committee on Health Research and Development of their "difficulty in obtaining adequate data from both governmental and non-governmental sources."[72] Hence this report recommended "that a system of health-related record linkage in Ontario, compatible with systems in other provinces, be introduced as soon as technical and legal measures have been devised to preserve the privacy of individuals."[73] The Council

70 Ibid., p. 28.

71 Ibid., pp. 18-19.

72 Ibid., p. 31.

73 Ibid., p. 66; see also p. 32.

added a particularly concise statement of the necessity for such a system of record linkages:

> "Studies of the etiology of disease and studies of the efficacy of preventive and therapeutic measures both require a capacity for bringing together records pertaining to one individual. This is particularly necessary in the study of the chronic diseases which make up a significant component of the total illness burden. In a chronic disease, etiological factors in the general environment or in the occupational environment preceed the onset of illness, often by many years. Similarly the benefits of preventive or therapeutic intervention may not be apparent until many years after their initial application. Hence the necessity for linking records of residence, employment and illness that have been made with respect to an indiviudal over a period of years and in different parts of the country."[74]

The report of the Council strongly supported the importance of promoting epidemiological studies in order to identify carcinogens derived from the environment.[75]

> "Epidemiological studies of cancer have identified clear-cut environmental factors of great causative significance. The direct relationship between cigarette smoking and cancer of the lungs is now accepted. Most recently, preliminary studies have shown relationships between other factors in the environment, including occupational hazards, and various forms of cancer."[76]

74 Ibid., p. 65.

75 Ibid., pp. 4, 39.

76 Ibid., p. 38.

The topics of record linkages and epidemiological research were also addressed in the October, 1977 report of the Science Council of Canada on Policies and Poisons.[77] This impressive report was produced by the Science Council Committee on Policies and Poisons, chaired by Dr. D. Bates, Dean of Medicine at the University of British Columbia. It examines specific knowledge about individual carcinogens, such as vinyl chloride, as a cause of occupational disease. This carcinogen causes a rare tumour of the liver among those employed in vinyl chloride plants. One of the lessons of discovering vinyl chloride as a cause of occupational disease is that "there is a need of methods of long-term follow-up if affected individuals who have left the industry are to be recognized 15 or 20 years after their exposures occurred."[78] In explaining why knowledge of dose-response relationships to vinyl chloride is incomplete, especially at low exposure levels, the report identified the lack of facilities for gathering pertinent information and the lack of systems for medical-occupational record linkage.[79]

77 Science Council of Canada, Policies and Poisons. The Containment of Long-term Hazards to Human Health in the Environment and in the Work Place. Science Council of Canada, Report No. 28 (Ottawa, October, 1977).

78 Ibid., p. 22.

79 Ibid., p. 23.

After a general review of knowledge concerning various carcinogens, the Science Council of Canada concluded that "the first task in the assessment of risk is the provision of the best possible information base. This necessarily includes: ... a study of morbidity statistics and analysis of record linkage and epidemiological information."[80] The Council concluded that:

> "There is an urgent need to design a medical record system in Canada that will permit the linkage of individual diagnosis with occupational and environmental history. We do not believe that such a system would represent any major infringement of privacy; its adoption is urgently indicated if the public is to be protected in the future against carcinogens which may only be manifest many years after the exposure experience has occurred .... We believe that the public would fully support such a record system in Canada, and we do not believe that medical confidentiality would be in any way compromised by its adoption."[81]

The Council recommended that "confident responsible persons who wish to conduct epidemiological research [should] have statutory right of access to the proposed medical record system." Although the Council advocated the initiation of such a record linkage system by the federal Department of National Health and Welfare, there is no reason that suitable beginnings cannot be made in Ontario.

80 Ibid., p. 34.

81 Ibid., p. 38.

CHAPTER IV

THE HOSPITAL MEDICAL RECORDS INSTITUTE (HMRI)

HMRI was established in 1963 under the sponsorship of the Ontario Hospital Association, the Ontario Medical Association, and the Ontario Hospital Records Association. At that time a similar type of system was in operation under the direction of Professional Activities Services, Ann Arbor, Michigan. A decade later the Ministry of Health required all hospitals in Ontario to enroll with HMRI. One immediate benefit of this action was to keep Ontario hospital data within Canada. HMRI obtained a federal charter in June, 1977 as a non-profit, non-commercial company. It has a fourteen-person board of directors nominated equally by the Ontario Medical Association and the Ontario Hospital Association. The directors include three physicians, three hospital administrators, two health record administrators, two representatives of the MOH, and four persons from the business community. HMRI is run on a daily basis by an executive director with the assistance of persons responsible for systems development, liaison services, production, and office administration.

HMRI is essentially a hospital patient data processing service located in Don Mills, Ontario. Between 275 and 300 separate

health institutions subscribe to the HMRI program today; less than ten are outside Ontario. HMRI's "intent is to provide professionals who are responsible for health care delivery, with an effective management tool of analytical and statistical utilization data to be used for planning, managing, monitoring trends, and evaluating patient care."[1] On January 1, 1978 HMRI introduced a revised reporting system developed by the Ministry of Health under the aegis of an HMRI Redesign Steering Committee. This was necessary to improve the turnaround time and contain costs.

HMRI receives data from hospital patients charts in the form of "basic abstracts" prepared by individual hospitals. Data are returned monthly and annually to the hospitals as both administrative and statistical reports. It is important to recognize that HMRI is not simply a statistical operation, although the HMRI literature states that its new system of reports is "designed to facilitate easy access to relevant, meaningful, statistical hospital data." But as will be discussed further below, the data returned to individual hospitals by HMRI can result in the identification of both patients and physicians. It would be possible for a hospital to take administrative action against individuals on the basis of

1 Hospital Medical Records Institute, The Hospital Medical Records Institute. A Health Data Processing Service (Don Mills, Ontario, 1978), p. 1.

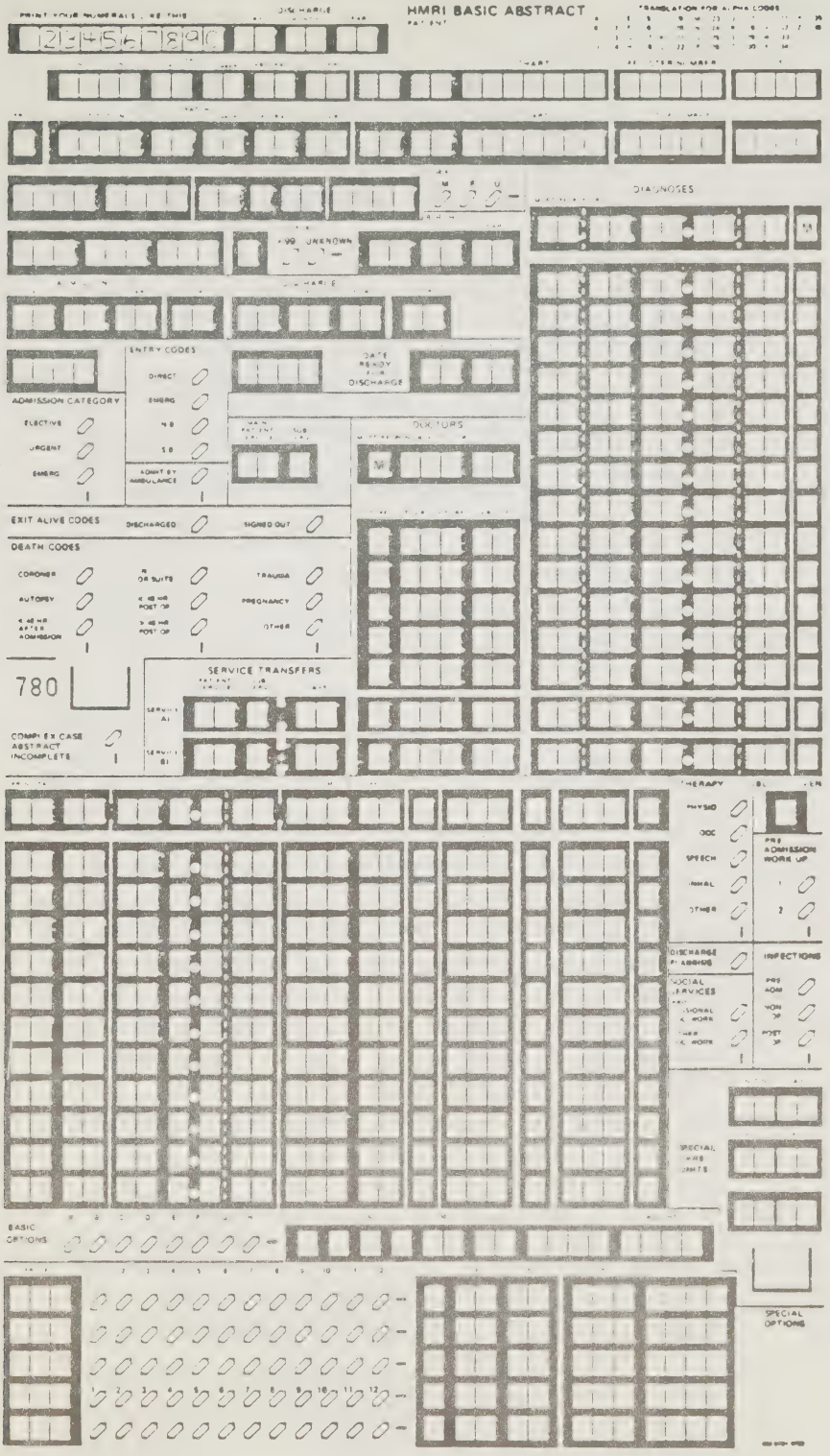
HMRI data. This issue raises a number of questions about confidentiality that are not really within the terms of reference of this Working Paper.

The main reason for including HMRI in this particular Working Paper is that it is heavily supported by the Ministry of Health of Ontario. HMRI now basically collects the hospital data for the Ministry of Health, as will be discussed further below. The first contract between HMRI and MOH was in 1975. By means of a contractual agreement dated April 1, 1978 between HMRI and the Minister of Health for Ontario, the Minister exercises substantial influence over the activities of HMRI. The contract specifies HMRI's operational services to Ontario hospitals and to the Ministry of Health, procedures for system evaluation and modification, accountability, and payments. The Minister of Health is represented in this connection by the executive director of the Information Systems Division or his designee. HMRI has to submit a monthly operating statement of budgets, income and expenditures, to the Ministry of Health. The HMRI systems for Ontario and the contents of reports cannot be changed without the approval of the Minister. MOH supports HMRI financially by establishing a fixed charge for each hospital abstract collected and provided to the hospitals, which pay HMRI directly. Thus although HMRI is an independent company and not a statutory governmental body, its contractual links highlight its financial dependence upon the Ministry of Health.

Data Collection

HMRI collects about 1.5 million hospital discharge records on individual patients in Ontario hospitals each year. The basic form used to collect microdata from each of 250 hospitals is called the "HMRI basic abstract;" a copy appears as Table X. In fact, the "abstract" is not much of an abstract, since it contains a significant amount of information. The abstract or input document for each patient separated from a hospital includes the following information: age and sex of the patient, admission and discharge information, diagnosis, procedures, hospital services, doctors and their role, and therapy and special care units. HMRI does not collect individual names on the grounds that these are not relevant to the basic uses to which hospitals put HMRI data; the abstracts are not identifiable by HMRI personnel. The hospital chart number permits the hospital to specifically identify a patient, whereas the OHIP number and sex and birthdate on the abstract would permit identification of an individual by the MOH. The OHIP number is furnished to the Ministry of Health, but a number of hospitals record that this information is not applicable or not available. The existence of the OHIP number, supplemented by the birthdate and sex of a patient, does provide a methodology for record linkages. In fact HMRI does not currently link data in any way. Nor does it have software in place with a capacity to search for specific individuals. Thus HMRI cannot search its machine-

TABLE X: HOSPITAL MEDICAL RECORDS INSTITUTE
BASIC PATIENT DATA ABSTRACT FORM



readable records to learn if a particular person has been treated in a hospital during a particular month.

Whenever a patient is discharged or separated from a specific hospital and his or her patient chart is completed, an abstract is prepared and sent to HMRI, where it is transferred to a magnetic tape by an optical scanner. The information on the tape is verified and validated and then reported to the hospital on a periodic basis.

It is evident that HMRI collects an extraordinary range of sickness-related data on individual patients and treatment-related data on individual physicians in a particular hospital. It returns the same data to the hospital in question in a variety of forms. Table XI is a listing of basic HMRI reports for each hospital. Several indexes and listings among these reports contain sensitive information. The monthly Diagnosis Index groups all patients assigned a specific diagnosis code. Each diagnosis lists the chart number of the patient, age, sex, length of stay in the hospital, and doctor and procedure information. The monthly Physicians' Index groups all patients with whom each physician was associated. The sample form used by HMRI to illustrate the monthly report named "index by physician" lists for a particular institution (identified only by a code) the specific diagnosis by code and descriptive title

TABLE XI: HOSPITAL MEDICAL RECORDS INSTITUTE
REPORTS FOR INDIVIDUAL HOSPITALS

HMRI REPORTS

TYPE	FREQUENCY
1. SUMMARY REPORTS	
Discharge Analysis—Part 1 & 2	Monthly & Annually
2. INDEXES	
Diagnosis	Monthly & Annually
Procedure	Monthly & Annually
Service	Monthly & Annually
Physician	Quarterly & Annually
3. LISTINGS	
Chart Number	Monthly
Chart Number Addendum	Monthly
4. LENGTH OF STAY REPORTS	
Hospital (Patient *&* Doctor Service)	Quarterly & Annually
Service (Patient *or* Doctor Service)	Quarterly & Annually
Diagnosis Groupings	Quarterly & Annually
Procedural Groupings	Quarterly & Annually
Physician (Most Responsible & Procedural)	Quarterly & Annually
5. SPECIAL REPORTS	
i.e., Accreditation Study O.R. Time Analysis Project Summaries and others.	

for a particular physician, who is identified only by a three-digit number. For example, a particular physician may be listed as having made one diagnosis for "drug dependence." The chart number of the patient is listed, the age of the patient, the length of stay in hospital, and various other diagnostic information.[2] The monthly Procedure Index groups all patients assigned a specific procedure code. The form used by HMRI to illustrate "index by procedure" lists for a specific institution a series of procedures identified by code and descriptive title. In some instances only one person has had a particular operation. The age, sex, and chart number of the person are given, along with a variety of data about doctors, procedures, and services.

In addition to the various indexes described above, HMRI Reports include a series of listings. The Monthly Chart Number Listing "displays virtually all the data from the abstract for each discharge. Data is sorted by ascending order of chart number."[3] This listing includes the chart number for each individual in a particular hospital, the birthdate, sex, doctor, diagnosis, and procedure information for each patient.

2 Ibid., p. 14.

3 Ibid., p. 15.

The contract between HMRI and the Ministry of Health defines the "Ontario Data File" as "the totality of data contained in the completed abstracts submitted to HMRI by all hospitals in respect of patients discharged since January 1, 1975 from such hospitals."[4] HMRI is required to supply MOH with all data accumulated in the Ontario Data File either for a specified period of time or on an annual basis. Attachment C, which is a part of the contract, specifies that the annual data from the Ontario Data File submitted to the Ministry "shall also contain complete patient identification as supplied by the hospital in respect of OHIP number, birthdate and sex, together with any unique individual identifier, such as the Ministry may require hospitals to supply in the future."

HMRI does not have its own computer. Data processing is performed in the Leaside Data Center run by the Ministry of Government Services (MGS). This is the same computer used by the Ministry of Health. HMRI argues that it needs a large computer for small amounts of time so that time-sharing is realistic. During 1978 HMRI will tender its computing job to a number of data processing utilities, including MGS, in part

4 Contract between HMRI and the Minister [illegible]lth for Ontario, April 1, 1978, s. 1.1(c). Herafter cited as Contract.

because it is sensitive to the issue of putting health data from another province into an Ontario government computer.

Confidentiality and Data Dissemination

As a matter of policy, HMRI "protects the confidentiality, privacy and security of all data referring to individual patients, professionals and participating organizations at all times."[5] The data collection and dissemination activities of HMRI are not covered directly by any Ontario statute. Through its involvement with specific hospitals, HMRI is indirectly subject to the Public Hospitals Act and its accompanying regulations. In agreement with that Act, HMRI regards the hospital as the owner of personal data and not the patients. Regulations under the Public Hospitals Act have not been issued to cover the inter-relationships among HMRI, individual hospitals, and the Ministry of Health.

HMRI does not have contracts with individual subscribing hospitals but with the Ministry of Health. This contract has a number of excellent provisions concerning confidentiality of

[5] The Hospital Medical Records Institute, p. 1.

personal data. The basic provision on the confidentiality of data is as follows:

> "8.1 Subject to anything to the contrary to this Agreement HMRI will at all times, both before and after the expiry or termination of this Agreement, preserve absolute secrecy with respect to any data in the Ontario Data File and, without limiting the generality of the foregoing, will
> (a) take all such precautions and measures as may be necessary to ensure that no person has access to any such data other than for the purpose of enabling HMRI to perform its obligations under this Agreement; and
> (b) not make any copies of any document, tape or other thing containing data from the Ontario Data File, except to the extent necessary for HMRI to perform its obligations under this Agreement."

The meaning of the first phrase in section 8.1 is unclear; the current language may contain typographical errors. The contract also requires HMRI to regulate its relationships with any subcontractors, who might obtain possession of any document, tape or thing containing data from the Ontario Data File, so as "to preserve absolute secrecy with respect to any such data and to take all such precautions as may be necessary to ensure that no employee of the subcontractor and no other person has access to any such data other than for the purpose of enabling the subcontractor to perform his obligations under his contract with HMRI."[6] HMRI does have subcontractors doing this type of work at present; from the point of view of data protection, it might be better if HMRI did not use subcontractors for data processing.

6 Contract, s. 6.3(a)(ii).

On the other hand, the use of subcontractors could be regarded as a positive benefit, since such an arrangement could prevent anyone employed by HMRI from having unrestricted access to the HMRI data files. Divided responsibility for security would at least require collusion between employees of different institutions before security could be readily breached. Another useful provision in the contract governing relationships between HMRI and any potential subcontractors requires HMRI to "take all such measures as may be necessary to retrieve any document, tape or other thing referred to above in the event the subcontractor does not return the same to HMRI forthwith after it has served the purpose of enabling the subcontractor to perform his obligations under this contract with HMRI."[7] This provision may help to prevent a type of situation that has arisen on occasion between subcontractors and the American federal government in the data collection field.

The contract also imposes restrictions on how the Ministry of Health may use data obtained from HMRI. In the first place "the Minister will not disclose any data taken from the Ontario Data File where he is prohibited by law from so doing."[8] Secondly,

7 Contract, s. 6.3(b).

8 Contract, s. 14.4.

> "Where the Minister proposes to use data derived from the Ontario Data File for any purpose other than epidemiological analysis and research, facilities and services planning, resource utilization and allocation, Program Evaluation or management of health care costs, the Ministry will consult with HMRI prior to implementing the analysis of the data."[9]

This provision expresses the desire of hospitals to remain well-informed about how the Ministry uses their data.

The contract between HMRI and MOH also regulates data dissemination for research purposes:

> "8.2 Notwithstanding paragraph 8.1, HMRI may disclose data to persons engaged in health care research, provided that
> (a) HMRI has first investigated and satisfied itself as to the merits of the research project and the integrity of persons engaged therein; and
> (b) in any event no data will be disclosed which might identify, or permit any person receiving such data to identify, any recipient or (unless the provider consents to such identification) any provider of health care services."

Thus no identifiable data can be released by HMRI without the consent of a hospital. Thus record linkages for research and statistical purposes, including epidemiological research, which normally depends on initial access to identifiable data, would require consent from a hospital. Another clause in the contract implicitly authorizes research uses of HMRI data. HMRI agrees that for each hospital it will "hold problem-oriented and other workshops for physicians, hospital administrators, coders and

9 Contract, s. 14.1.

researchers for education and training in the use of the system."[10]

It seems obvious that the various types of HMRI reports to specific hospitals, including summary reports, indexes, listings, length of stay reports, and special reports, can have considerable utility for research and statistical purposes, in addition to their obvious administrative applications. Table X also indicates that the HMRI basic abstract form can be used to collect and code the data for a specific research project. HMRI illustrative material describes this capacity as follows: "The capture and reporting of Medical Audit information is facilitated in the redesigned system by means of a "project area" which will be used mainly for retrospective studies. When the Medical Audit Committee within a hospital has established criteria to be used for a specific study, it becomes a simple matter for the record staff to imput this information as they abstract the record." The area on a coding sheet could thus be used for a traditional research project or for an administrative study such as the monitoring of patient care by a particular physician.

HMRI has the capacity and the willingness to be of considerable assistance in the use of hospital patient data for research and statistical purposes. This remains true, despite the fact that

10 Contract, s. 2.1(g).

research and statistical uses would be considerably subordinate to the overwhelmingly administrative tasks of HMRI data files. Requests for access to data for scientific purposes can be sent directly to HMRI, which is prepared to service users directly with hospital data. HMRI only releases data for an identifiable institution after it has received authorization from the originating institution to make the data available to a researcher. Although the intervention of the MOH is not necessary to secure access to hospital-related data, it would be normal in a research program to combine hospital and medical data, thus necessitating the involvement of the Ministry. One of the major burdens in using HMRI data for such purposes is the cost of accessing and reviewing large amounts of hospital data for a great many hospitals or on a province-wide basis. Thus cost and not the lack of utility of the data tends to discourage many users, such as graduate students, who cannot afford the cost of computer time. HMRI is studying the patterns of ad hoc requests for access to data for research and statistical purposes, so that it might be possible to spread the costs among several potential users over a period of time.

In the last eighteen months HMRI has helped one research project. It assisted the Health Care Research Unit of the University of Toronto, which is conducting a major study for the Ministry of Health of trauma-related hospitalization. HMRI obtained release

forms for the hospital patient data from the hospitals and not from the individual patients. The researchers wanted data elements for twelve Toronto-area hospitals where patients were admitted during 1976 with trauma-related diagnoses. HMRI furnished the researchers with individual data on the number of the institution, the chart number and register number of the patient on admission, the birthdate, sex, ambulance number, admission and discharge dates, and diagnostic information. The researchers received this data from HMRI as microdata. The Ministry of Health furnished the researchers with even more data, and hospitals gave them access to patient charts.

Recommendations

The existence of HMRI should facilitate the use of Ontario hospital data for legitimate research and statistical purposes. HMRI and the Ministry of Health should promote such goals in the interests of the general public. Various mechanisms, such as advisory committees, should be created so that research projects can be approved and reviewed in a professional manner. HMRI already collects a limited amount of identifying information about each patient discharged by the hospital. The adoption of a Unique Personal Identifier will render the collection of named data unnecessary. If, as seems evident, Unique Personal

Identifiers are ultimately essential for epidemiological research, then the HMRI basic abstract should begin to include them. HMRI should be encouraged to examine the pattern of ad hoc requests by research users, so that the costs of servicing researchers can be reduced to a tolerable level. HMRI should also be in an excellent position to produce public use samples for general research and statistical purposes.

CHAPTER V

ONTARIO CANCER TREATMENT AND RESEARCH FOUNDATION (OCTRF)

The Ontario Cancer Treatment and Research Foundation was incorporated in 1943 as a foundation by the Ontario legislature and reconstituted in 1957 by the Cancer Act.

> "The object of the Foundation is to establish and conduct a program of research, diagnosis and treatment of cancer, including, ...
> (f) the adequate reporting of cases and the recording and compilation of data; ...
> (h) the providing of facilities for undergraduate and post-graduate study; ...
> (j) the providing and awarding of research fellowships....[1]

The Foundation thus has research, diagnostic, and treatment functions. OCTRF administers and finances six Regional Treatment Centres for cancer in Ottawa (two divisions), Kingston, Hamilton, London, Windsor, and Thunder Bay. It also partly funds research at the Ontario Cancer Institute, which incorporates the Princess Margaret Hospital in Toronto, and research at other centres.

The funded research includes laboratory, clinical, and epidemiological investigations, including environmental

1 Ontario Cancer Act, R.S.O., 1970, c. 55, as amended by S.O. 1972, c. 34, s. 5.

carcinogens. Through its Division of Epidemiology and Statistics the Foundation conducts significant research in the field of biostatistics. For forty years from 1936 to 1976 Dr. A.H. Sellers was associated with the Ministry of Health and OCTRF in the areas of medical records and the development of statistical-epidemiological research. He was responsible for the development of most of the practices described in these pages.

Sources of Data

OCTRF continues to acquire and store cancer-related personal data. It receives from the Ministry of Health hospital separation froms or the equivalent with mention of cancer as a discharge diagnosis. The death abstract service furnished by the Office of the Registrar General has already been discussed in Chapter I.

OCTRF appears to have three major data bases that are relevant to the concerns of this Working Paper.[2] The Cancer Death File contains data from medical certificates of cause of death for all

1 The descriptions that follow are derived from interviews and from the "Brief to the Royal Commission of Inquiry into the Confidentiality of Health Records in Ontario," submitted by the Ontario Cancer Treatment and Research Foundation (June, 1978), pp. 1-2, 15-19. Hereafter cited as OCTRF, Brief.

persons dying in Ontario from 1953 to 1977 with or from cancer, including age, sex, residence, and site of the cancer. The file currently covers 175,000 deaths identified by name, date of death, and serial number of death registration. The data are stored on punch cards and computer tapes for an indefinite period. In terms of particular research, the Cancer Death File can be used to confirm the death of a patient, whose whereabouts have been lost track of; the death can then be followed-up by the Foundation and other clinics and registries for research purposes. This file can also be used for occasional special studies.

A second major data file is the Ontario Cancer Incidence Survey and Registry, which contains information on the age, sex, residence, site, extent, histological type, treatment, and subsequent course for 160,000 cancer patients. The file began as a pilot study of the incidence of cancer in Ontario from 1964 to 1966 and is being continued as a basis for a provincial cancer incidence registry in order to provide data on newly-diagnosed cancer cases throughout the province. The sources of approximately 700,000 reports on episodes of cancer are the Regional Treatment Centres, clinic-associated and other hospital registries, pathology reports, OHIP hospital separation forms, drug prescription forms, and death registrations. OCTRF has data

in this file for 1967 through 1976 from all the sources listed.[3] Individual cases are uniquely identified by such items as name, age, sex, address, birthdate, and death date. The data are stored for an indefinite period in the form of index cards, computer tapes, and listings. Cases from 1969 to 1971 are now linked for the production of incidence statistics. This survey file is a source of data to help meet the operational and planning needs of the Foundation.[4]

The third significant data base at OCTRF is the New Case File based on reports of cancer from the Regional Treatment Centres and associated registries. The New Case File contains data on patients newly-admitted to the Foundation's clinics, with name, age, sex, dates of registration, previous clinic treatment site, extent, and histological type of cancer, and eventually follow-up data. The File contains 102,000 cases registered from 1960 to 1976. The system pertains primarily to patients referred for radiotherapy, and includes approximately fifty percent of all patients with cancer in Ontario: "The data are in machine-readable form from 1960 to 1977; from 1936 they are available on abstract cards."[5] The intent of the File is to provide

3 Ontario Cancer Treatment and Research Foundation, Cancer in Ontario, 1977 (Toronto, 1977), p. 71.

4 Ontario Cancer Treatment and Research Foundation, Cancer in Ontario, 1976 (Toronto, 1976), p. 72.

5 OCTRF, Brief, p. 1.

Foundation clinics with the essential details about their new cases, eventually with survival rates, and to provide information on aspects of cancer in response to any inquiry from the Foundation and the public. The uniform case recording and follow-up system was established as long ago as 1937 to record new cancer cases registered at the Regional Treatment Centres. The computer-based cancer incidence and registry program now summarizes the clinical course of over 225,000 cancer patients; about 40,000 have been reviewed in detail for OCTRF's clinical conferences.[6] Two teaching hospitals in Toronto operate approved cancer registries and report their findings to the Foundation.

Policy on Confidentiality and Data Dissemination

The June, 1978 brief of OCTRF to the Royal Commission on Confidentiality of Health Records highlighted the strong commitment of the Foundation to protecting the confidentiality of data entrusted to it. At the same time the brief recognizes that identifiable data or at least data with pertinent identifiers are essential for the maintenance of cancer registries. As the late medical director of OCTRF commented in 1977 in connection with the Foundation's medical records and statistics,

6 OCTRF, Cancer in Ontario, 1976, pp. 59-71.

"it is an unending task to assure that the records kept are completely accurate, and the statistics derived from them are dependable. Every precaution is taken to maintain complete confidentiality as far as individual patients are concerned."[7]

From a purely legal perspective, the Foundation is fortunate in having a strong statutory statement on confidentiality, which was added to the Cancer Act in the Cancer Amendment Act of 1972: "Any information or report respecting a case of cancer furnished to the Foundation by any person shall be kept confidential and shall not be used or disclosed by the Foundation to any person for any reason other than compiling statistics or carrying out medical or epidemiological research."[8] Unfortunately, the clause does not specify any sanctions. The Regional Treatment Centres and the Princess Margaret Hospital also operate under the provisions of the Public Hospitals Act and regulation 729, which have already been discussed in Chapter III, from the point of view of confidentiality and data dissemination. In addition, all members of the Epidemiology and Statistics Division swear an oath of secrecy before the Foundation's solicitor; the oath can be reviewed in Table XII.

7 OCTRF, Cancer in Ontario, 1977, p. 15.

8 Cancer Amendment Act, S.O. 1972, c. 34, s. 6a (1).

TABLE XII: OATH OF SECRECY FOR SELECTED STAFF

THE ONTARIO CANCER TREATMENT AND RESEARCH FOUNDATION

OATH OF SECRECY

I, ...
(given names)

.....
(surname)

do solemnly swear that I will hold secret and will not disclose or communicate to any person any information or document given me from the records of the Foundation or obtained from those records by reason of my access thereto, except as required in the performance of the duties of my office.

SWORN before me at the
of
in the of .....
(Signature of Deponent)
this
day of 19

.....
A Commissioner, or a
Notary Public.

In his 1977 report on Ontario cancer statistics, Dr. Sellers pointed out the existing controls on access to and use of the Cancer Registration and Activity File:

> "All Centres should have access to the statistical data in the Registration File, provided that individual patients or Centres cannot be identified; if identification is possible, or the data are sensitive, permission must first be obtained from the directors of the Centres involved .... When the new computer system ... is in operation, a subcommittee will be established with one representative from each Centre, to be responsible for the facilitation, supervision, and control of access to the computer-based patient files.[9]

The establishment in November, 1977 of a Subcommittee on Access to and Use of Patient Files has been one response to OCTRF's concern over the issue of confidentiality and privacy. It was argued that a visible, formal procedure with numerous safeguards was required to prevent unauthorized access to the Foundation's records. It is a Subcommittee of the Standing Committee on Records and Statistics and is composed of the Medical Director of the Foundation, the Head of the Division of Epidemiology and Statistics, the Systems Manager of the Computer Centre, and the Chairman of the Medical Records Committee of Princess Margaret Hospital.[10] One of the tasks of the Subcommittee is to handle requests for information on patients in the Regional Treatment Centres. This Subcommittee can also

9 OCTRF, Cancer in Ontario, 1977, p. 80.

10 OCTRF, Brief, p. 4.

review dissemination to outsiders of any data collected or stored by OCTRF. The Foundation's brief to the Commission on Confidentiality of Health Records points out that in the past "there have been two requests from researchers in the cancer field for data with identifying information, specifically name. It has been possible for the staff at Head Office to process these requests and provide the researchers with the required data in such a form that no identification of individuals was possible."[11] If requests cannot be satisfied in this manner in future, they will be referred to the Subcommittee, which will screen all requests for access to identifiable information. Table XIII contains an oath of secrecy for outside persons granted access to OCTRF records for research and statistical purposes approved by the Foundation.

The brief to the Commission on Confidentiality of Health Records also pointed out the numerous procedures that OCTRF has instituted in order to protect the confidentiality of data.[12] Data from source documents furnished to the Foundation are coded by the sworn staff of the Epidemiology and Statistics Division. All documents containing identifiable information are then shredded on the premises of the Foundation. Since the Foundation

11 OCTRF, Brief, p. 4.

12 OCTRF, Brief, p. 46.

TABLE XIII: OATH OF SECRECY FOR RESEARCHERS
ONTARIO CANCER TREATMENT AND RESEARCH FOUNDATION

IN THE MATTER OF THE CANCER ACT, R. S. O. 1970
CHAPTER 55, AND AMENDMENTS THERETO

I, ...

being granted access to the records of The Ontario Cancer Treatment and Research Foundation for the purposes of compiling statistics and carrying out medical and epidemiological research approved by the Foundation,

DO SOLEMNLY SWEAR

that I will hold secret and will not communicate or disclose to any unauthorized person, any information coming to my knowledge, or the contents of any document given me from the records of the Foundation or obtained from the records of the Foundation by reason of my access thereto, except such statistics, or the results of such research.

SWORN before me at the)
of in the)
of)
this)
day of 1971.)
)
A Commissioners, etc.)

(Signature of Deponent)

now has its own dedicated computer, it is intended that data will be entered directly on the computer by the Foundation's personnel. Personnel from outside the Epidemiology and Statistics Division do not have access to identifying documents, which are generally kept under lock and key. Measures also exist to protect the confidentiality of documents in transit to the Foundation. There are other security measures proposed to control access to the Foundation's computerized data bases. There will be dedicated phone lines between the Foundation and the Regional Treatment Centres, user control numbers, audit trails of attempts to enter the computer system, and burglar alarms for after hours.

Uses of OCTRF Data

OCTRF is not simply content to record and compile data:

> "The data must then be used to study both the distribution and determinants (epidemiology) of cancer in Ontario. Epidemiological studies can thus provide hypotheses as to the etiology of cancer. It has been estimated that environmental factors contribute to the production of the majority of cancer; and that, if these factors could be identified, it is possible that almost eighty percent of cancers could be prevented. Prevention of disease is almost always less expensive than the treatment of established disease."[13]

13 OCTRF, Brief, p. 1.

The Brief further emphasizes that "registry records enable researchers to identify trends in cancer incidence; these data are a critical supplement to mortality trends, which are less easily interpreted because mortality is affected by treatment. Incidence trends therefore provide more reliable grounds for research on new etiological agents. An active cancer registry is a vital tool for epidemiologists and other workers in the health field. It can assist materially in the search for occupational and environmental carcinogens, the investigation of diseases in individuals and families, and the identification of high-risk groups."[14]

Since the annual volumes published by OCTRF on Cancer in Ontario describe in considerable detail the research projects undertaken by the Foundation and its research associates, it is unnecessary to describe this work in detail here. Some of the research projects described above in connection with the Office of the Registrar General and the Ministry of Health by such researchers as Dr. Aileen Clarke and Professor David Hewitt have depended in part on access to data at OCTRF. Dr. Clarke, who is also now head of the Division of Epidemiology and Statistics at the Foundation, has also conducted internally a linkage study

14 OCTRF, Brief, p. 2.

between cancer mortality and the incidence file of osteogenic sarcoma to study survival rates from changing forms of treatment of this disease. Table XIV illustrates a recent effort by the Foundation to collect new information concerning the number of new cases of acute lymphatic leukemia diagnosed between 1970 and 1977.

Recommendations

OCTRF appears to be an organization that has made considerable and successful efforts to put its own house in order in connection with the protection of confidentiality and the promotion of legitimate research. Public hospitals, the Ministry of Health, and the Office of the Registrar General should continue to furnish cancer-related data to the Foundation. Dr. Sellers recently pointed out that "access to hospital separation forms with mention of cancer (or their equivalent) is essential to Ontario cancer incidence and registry program development, and the Foundation has been assured that the present information-sharing arrangements with the Ministry of Health will continue."[15] It is reported that a few Ontario hospitals are still reluctant to furnish the Foundation with copies of pathology

15 OCTRF, Cancer in Ontario, 1977, p. 72.

The Ontario Cancer Treatment and Research Foundation

7 OVERLEA BOULEVARD
TORONTO, ONTARIO M4H 1A8
TELEPHONE: (416) 423-4240

April 11, 1978

The Division of Epidemiology and Statistics of the Ontario Cancer Treatment and Research Foundation has received opinions from several sources in the last few months about a possible, recent increase in the incidence, in Southern Ontario, of acute lymphatic leukemia (204.0). The question has been raised as to whether this might be related to particular carcinogens or environmental agents.

The first priority is to determine whether the incidence of acute lymphatic leukemia has in fact changed. The Foundation has information regarding the incidence of leukemia between 1966 and 1970, and in order to determine whether there has indeed been an increase in the incidence since that time, I would like to enlist your cooperation to determine the number of new cases of leukemia diagnosed between January 1, 1970 and December 31, 1977. The information required to ensure that no single patient is counted more than once includes the name, address, date of birth, diagnosis, and date and place of diagnosis.

Under the amendment to the Cancer Act, a copy of which is enclosed, you will note that such information furnished to the Foundation is kept confidential and that the amendment protects physicians against liability in furnishing the information.

In view of the agents suggested as possible causes of the increase, I would appreciate having this information forwarded to me, as soon as possible, by registered mail to the above address. Should you wish to use the attached form to forward the information, please feel free to do so.

If an employee of the Foundation, who has sworn our oath of confidentiality, would be able to assist your staff in obtaining this information, I would be glad to arrange this upon hearing from you. Should you have any questions, please do not hesitate to contact me at 423-4240.

Thank you for your help.

Sincerely yours,

E. Aileen Clarke, M.B., B.S., M.Sc.,
Head,
Division of Epidemiology and Statistics.

Encls.

reports of cancer for fear of re-contact with patients by researchers. This concern seems ill-founded, since the Foundation has careful procedures for screening researchers. A proposal must first be approved by an OCTRF committee and then by a particular hospital before a researcher would even approach a physician for permission to re-contact a patient.

Data on hospital patients collected by HMRI is currently of no use to OCTRF, because HMRI does not collect identifiable data. Perhaps this situation will change after the introduction of Unique Personal Identifiers and their collection by HMRI. The Foundation has substantial need for Unique Personal Identifiers because of the difficulty of linking records in the provincial Cancer Incidence Registration File. Dr. Sellers has pointed out that "because no unique identifying number is available, machine linkage of the records has been a formidable project, but successful linkage was achieved and well tested by October, 1975. The file was then merged to provide a master file sorted by site, year of birth, and sex."[16] By August 1, 1977 the necessary visual editing had been done; requests for existing information are frequently received.

16 Ibid., pp. 71-72.

The Brief of OCTRF to the Ontario Commission on Freedom of Information and Individual Privacy advocated the creation of an Ontario Task Force on Confidentiality of Computerized Medical Records. This suggestion is derived from the model of a New York State Committee that formulated ethical guidelines for centres handling medical data.[17] Since two Ontario Commissions are currently looking into matters of confidentiality of health records, this suggestion may not be necessary. Given the description above, it should be evident that OCTRF has already introduced many of the necessary protections for the confidentiality of data. There is also the risk of Ontario attempting to reinvent the wheel, given the existing literature on this sensitive topic from countries such as the United States and Great Britain.

Both the Medical Research Council and the Royal College of Psychiatrists in Great Britain have developed significant documents concerning the use of identifiable data in research.[18] The material prepared by the Royal College of Psychiatrists is particularly significant because it deals with the collection of

17 OCTRF, Brief, Appendix D.

18 Medical Research Council, "Responsibility in the Use of Medical Information for Research," British Medical Journal, (January 27, 1973), I, pp. 213-16; Royal College of Psychiatrists, "Confidentiality of Psychiatric Data in Medical Information Systems," British Journal of Psychiatry, CXXVIII (1976), 417-27.

identifiable personal data on psychiatric patients in public hospitals for the government's Mental Health Inquiry.[19] The document in question contains a sophisticated discussion as well of technical arrangements for insuring confidentiality. It would also be useful for OCTRF and comparable custodians of sensitive health-related data to review the recent memorandum from the Scottish Home and Health Department, dated March 20, 1978, concerning the confidentiality of medical records held on computers. These include rules governing the release of data and guidance on physical security of computers and computer-held data. The British have recently devoted considerable efforts to this area of data protection and data dissemination because of the extensive hearings held by the government's Data Protection Committee; its report is expected to be published within several months.

19 See David H. Flaherty, Privacy and Government Data Banks. An International Perspective (New York, 1978, forthcoming), chapter 3.

CHAPTER VI

THE ALCOHOLISM AND DRUG ADDICTION RESEARCH FOUNDATION

The Alcoholism and Drug Addiction Research Foundation is "an agency of the province of Ontario which operates specialized research, educational, clinical, and service development programs throughout the province."[1] It is a crown corporation governed by a statute known as the Alcoholism and Drug Addiction Research Foundation Act of 1965.[2] Overall policies are set by members of the Foundation, who are community representatives appointed by the Lieutenant Governor in Council. ARF also has a professional advisory board. The provincial headquarters of the Foundation are in Toronto; in addition, there are 34 regional offices.

ARF is of considerable interest from the point of view of this Working Paper, because it is a government agency as well as a treatment and research foundation. According to its governing statute, the Foundation has the power:

1 Alcoholism and Drug Addiction Research Foundation, Annual Report, 1976-77 (Toronto, 1977), p. ix.

2 Alcoholism and Drug Addiction Research Foundation Act, 1965, R.S.O., 1970, c. 18.

> "a) to conduct and promote a programme of research in alcoholism and addiction; and
> b) to conduct, direct and promote programmes for, ...
> iv) the dissemination of information respecting the recognition, prevention and treatment of alcoholism and addiction.[3]

Two of its major entities are a Clinical Institute Division and a Research Division. There is also a division handling regional programs. The Clinical Institute is a hospital designed to provide a facility for treatment research and clinical investigation. The Research Division has six sections specializing in various kinds of studies. The sections focusing on Psychological Studies, Social Studies, Evaluation Studies, and Research Planning, including a Statistical Information Unit, are most likely to use ARF and other government data. There appears to be a consensus within the Foundation that the principal function of ARF is to carry out research, and that all treatment carried out by ARF is undertaken primarily for therapeutic research purposes.

Sources of Data

ARF collects substantial amounts of personal data on its own patients (administrative data) and on survey populations for research and statistical purposes. The data are particularly

3 R.S.O., 1970, c. 18, s. 7.

sensitive because they deal with the population of alcoholics and drug addicts.

The ARF data base known as "patient records" contains information on 10,000 patients identified by patient number. The data are stored as manual files and on computer tape and are strictly confidential. The contents include "demographic information on patients and patient-related events, i.e., admission, interviews, discharge."[4] The purpose of "patient records" is "to provide data on the characteristics of patients and statistics such as number of interviews, source of referral."[5]

The second major data set known as "various ad hoc research surveys" is maintained by the Research Division of ARF. The range of data includes surveys on the prevalence of alcoholism, chronic drunkenness, alcohol buying habits, alcohol and traffic accidents, medical prescription drugs, the mortality of alcoholics, and the non-medical use of drugs by secondary school students.[6] The surveys are identified by project name or

4 Ministry of Treasury, Economics and Intergovernmental Affairs, Catalogue of Statistical Files in the Ontario Government 1977 (Toronto, 1977), p. HL36. Hereafter cited as Catalogue.

5 Catalogue, p. HL36.

6 Catalogue, p. HL1.

characteristic and average about 6,000 individuals in size. They are stored as paper files, punch cards, computer tapes, and photocopies. The stored survey data contain aggregated, non-identifiable data. The objective of these various ad hoc research surveys is "to provide data for epidemiological research designed to map the extent, quantity, frequency, load and consequences of the use of alcohol and other psychoactive drugs in Ontario."[7] The Catalogue of government data files has the peculiar notation that the surveys are "not confidential," except that where the data base is a medical record the usual rule of confidentiality applies. In fact, all respondents to ARF surveys receive promises of confidentiality.

The Research Division of ARF evidently collects a great deal of their own data for research purposes, as a review of research project descriptions in the Annual Report of the Foundation indicates.[8] For example, one large sample used for research purposes is a collection of primary data on the mortality experience of alcoholics, mainly from the Toronto area. It comprises a sample in excess of 100,000 man years. Data have also been collected for a study of 3,500 acute drug ingestion

7 Catalogue, p. HL1.

8 ARF, Annual Report, 1976-77, pp. 14-18.

and drug abuse patients in 21 emergency rooms of metropolitan Toronto hospitals. A recent survey in the London area collected current information on the number of alcohol and drug abusers receiving care, and the nature and extent of the services provided. There is also a major longitudinal study of marijuana use among Ontario high school students.

ARF also has access to data from the Office of the Registrar General and the Ministry of Health. One ARF researcher, Wolfgang Schmidt, uses mortality data from the Registrar General for morbidity studies. The Data Development and Evaluation Branch of MOH furnished the ARF Task Force on Treatment Services for Alcoholics with tabulations of hospital in-patient data on a provincial basis, for specific regions, for specific diagnoses and by county of residence of the patient. DDEB used the hospital in-patient discharge file and the Ontario Psychiatric File in order to produce these special tabulations. The Ministry would only do a limited number of tabulations of this nature because of the burden of work involved. Some of the research work on detoxification conducted by ARF has also involved access to MOH data.

Policy on Confidentiality and Data Dissemination

ARF has recently gone to considerable lengths to develop sound policies on the protection of confidentiality and the regulation of data dissemination for research and statistical purposes. In a letter dated August 12, 1977, that was submitted to the Commission on Freedom of Information and Individual Privacy, the President of ARF, John B. Macdonald, made the following statement:

> "ARF is acutely aware of its responsibility to protect its patients' right to the maintenance of their records in a confidential manner (for their alcohol-and-drug related problems carry considerable social stigma even in the 1970's). However, because of our extensive research commitment, ARF is also vitally concerned that investigators have access to whatever patient records are necessary to adequately meet their research objective. Moreover, pursuit of some lines of research by ARF investigators requires access to patient and/or statistical information from the Ministry [of Health] (e.g. from HMRI, Drug Benefit Program records, etc.) relating to persons who are not ARF patients."

This statement accurately reflects the dilemma that ARF has to resolve as both a custodian of data and a substantial research user. In the same letter the President of the Foundation also reports that any research investigator or their support staff are expected "to observe the same standards of confidentiality concerning information to which he/she has access via the patient record as are physicians. Breach of confidentiality (intentionally or through negligence), by any such person constitute grounds for disciplinary action."

Table XV is the most general section of the six-page ARF "Policy on Confidentiality of Patient Information." Sections A and E of this general statement are particularly relevant to the concerns of this Working Paper. Because the Clinical Institute of ARF is essentially a hospital, it is subject to the Public Hospitals Act and Regulation 729, especially subsection 48, on access to hospital patient data. This statute and the regulations control access to patient data within ARF and also determine what data ARF has to furnish to the Ministry of Health and HMRI. The same situation prevails for the Princess Margaret Hospital of the Ontario Cancer Institute, which was mentioned above in connection with the Ontario Cancer Treatment and Research Foundation. The confidentiality provisions of the Public Hospitals Act and the accompanying regulations are particularly useful for ARF because the governing statute of the Foundation does not provide protection for the confidentiality of data. The general consent form for ARF patients mentioned in Section E of Table XV is reproduced here as Table XVI.

ARF also subscribes to the University of Toronto's "Guidelines for the Use of Hospital Records for Research Purposes," which were prepared in 1977 by the Human Experimentation Subcommittee of the Research Board of the University of Toronto. This excellent document, which was adopted in principle by ARF, is included here as Table XVII.

TABLE XV: POLICY ON CONFIDENTIALITY
OF PATIENT INFORMATION
ALCOHOLISM AND DRUG ADDICTION
RESEARCH FOUNDATION

The following is the policy of the Addiction Research Foundation on the release of patient records or patient related information. Information in patient files or any information that has been received from the patient, documented or not, is not to be made available except:

a) where the person seeking patient information is specifically covered by Regulations of the Public Hospital's Act, Section 729, subsection 48 (copy enclosed).

b) when, in the interests of the patient, and with his/her written consent, we provide information for use in a court case.

c) when, in a court case the patient record is subpoenaed. The file is then delivered by a member of our patient record staff and/or therapist who will then interpret the contents of the file to the court but only after direct request by the presiding judge to reveal the contents of the file.

d) in a criminal investigation when a search warrant is provided by the investigating officers, then patient information must be made available as requested by the warrant.

e) when the patient has signed the general consent on the Face Sheet permitting use of files and information by our Research staff for research purposes*.

*Information and records which existed prior to the introduction of the Face Sheet consent form will fall outside of this regulation when the information is used for ethically approved research projects.

TABLE XVI: PATIENT CONSENT FORM
ALCOHOLISM AND DRUG ADDICTION
RESEARCH FOUNDATION

I understand and agree:

1. That my medical records may be examined for research purposes by the Addiction Research Foundation staff.

2. That I may be approached by trainees and be asked to volunteer my cooperation for teaching exercises.

3. That the Foundation may wish to contact me at a future date to obtain follow-up information.

I further understand that the Foundation will maintain confidentiality in the manner in which medical records are usually treated.

Note: Failure to sign this consent form will not result in any changes to the treatment program.

_____ _____

Witness Signature

Source: Face Sheet-Patient Data Form (reverse side). Clinical Institute. Alcoholism and Drug Addiction Research Foundation.

TABLE XVII: UNIVERSITY OF TORONTO GUIDELINES FOR THE USE OF HOSPITAL RECORDS FOR RESEARCH PURPOSES

The following report of the committee on the use of hospital records for research purposes was accepted March 29 by the human experimentation subcommittee of the University's Research Board *(Ref. Bulletin, April 15.)*

Terms of reference

"In view of the wide differences that exist among the review committees in the requirements made of investigators with respect to access to hospital records, the human experimentation subcommittee has been asked by ORA to consider ways in which it might guide the thinking of review committees and offer useful guidance to deparment chairmen whose responsibility it is to make known to members of their departments the University's requirements with respect to research conducted under its auspices."

The terms of reference and the complex nature of the subject suggest that the committee should seek to enunciate general principles and essential considerations rather than detailed and highly specific rules.

For our purposes we considered hospital records to be all information concerning patients retained as:
(a) Statistical material;
(b) Case histories filed in conventional ways;
(c) Data banks.

The purpose of recording and retaining information is to provide the basis for diagnosis and treatment of the patients whose records they are, and for increasing medical knowledge, through research. The ultimate goal of research is to increase the accuracy of diagnosis and effectiveness of treatment for all members of society.

The ethical principles that are involved in the use of hospital records concern the rights of individual patients and the needs and rights of society. In an ultimate sense these may coincide, but in some situations they appear to be in conflict. The individual has the right to:
(a) **privacy** – achieved through confidentiality in the use of information about him, and discretion in exposing his person in the presence of others;
(b) **self-determination** – achieved through the exercise of choice and consent;
(c) **respect** – achieved through courtesy and sensitivity to his feelings.

What we refer to as the rights of society in this context are the rights of everyone to the best care that medical knowledge makes available, and consequently to the most advanced knowledge that the profession can obtain.

The ethical responsibility of all concerned is therefore to balance the need for accessibility of patient information for research purposes with the obligation to respect the individual patient's rights as stated.

Access to records

If hospital records are to be accessible for research purposes, the interests of privacy must be served by the restrictions that are imposed upon the accessibility – governing the purpose for which they are used and the persons who may use them. The purpose should be for research of high quality aimed at the public good. Those having access should be qualified professionally, and directly associated with the patient's care and/or with the institution having custody of the records; or they may be government employees directly responsible for society's health care; or they may be qualified research personnel from institutions such as universities, which have accepted standards of protocol and established arrangements with hospitals concerning the purpose and procedures of specific research activities.

The majority of reputable hospitals have review boards with authority to assess and approve the use of patient records for research purposes by members of the hospital staff or of associated institutions, and to ensure that there are adequate safeguards to protect the confidentiality of the information. However, of increasing concern is the question of the extent to which the use of records should depend upon the consent of the patients concerned, in principle and in practice. The urgency of this matter is increased by the establishment of data banks with computer-facilitated accessibility from the remotest regions by radio and telephone. To the practical difficulty of obtaining consent for retrospective studies in the unforseeable future is added the complexity of monitoring the use of data banks.

Patient consent

General consent

There is a kind of consent on the part of the patient that is implicit when he

TABLE XVII (cont'd)

understands and accepts (a) the connection between his expectations of the medical profession and the availability of his medical records for research purposes, and (b) the responsibility of government for public health, and its need for statistical and other information which he can help to provide. This kind of understanding and consent is more a matter of *education* than of formal statement, though there are many situations where such a statement would be appropriate.

An extension of this is a kind of general consent that is implicit when a patient entering hospital understands and accepts normal hospital practices that involve the use of records for statistical purposes and research. These practices and the reasons for them should be explained in appropriate literature made available on admission. Information should also be given about affiliations with other hospitals, with universities and with other research institutions. Reference should be made to the nature and importance of retrospective studies, statistical studies, team research and individual research. The safeguards which protect the confidentiality of information should be carefully described. A statement of the relevant codes of professional ethics should also be given.

With this information clearly presented, it should be reasonable to expect that explicit and particular consent is not necessary in the case of statistical studies by authorized persons following the approved procedures and supervised by the hospital authorities. By the same token, and under the same controls, explicit and particular consent should not be necessary for a physician to use his patients' records for his own research purposes, with a view to publication, or to make his patients' records available to other physicians for the same purpose, or to collaborate with approved researchers from universities for a similar purpose. Generally, the permission of the physician should be sought by anyone wishing access to his patients' records, and only medical personnel or those approved by the hospital board should be given this permission.

Particular consent

When a research project proposes a use of records which makes the risk of identification unusually great, or otherwise intrudes upon the patient's privacy or that of his family, particular and explicit consent should always be sought - by his own physician or any other authorized person.

Consent and coercion

There is an inherently coercive element in a situation in which a patient requires medical care and is asked to consent to such things as the use of his records for general, albeit research, purposes. In a sense, the patient is at a disadvantage. Furthermore, the more the information tends to reveal his identity and the less directly the benefits appear to accrue to him, the more difficult it is for him to see the justification for giving consent. The tension of this situation is unavoidable, but it can be reduced significantly by keeping the patient well informed and by treating him as a collaborator in the enterprise of research, rather than as a victim - a passive instrument of what is often seen to be someone else's professional ambition.

The University's responsibility

Academic research institutions have a responsibility to educate, as well as to do research. They have also an ethical responsibility to protect the values of society in the process of research.

The University should, therefore:

(a) publicize the general principles governing professional, social and ethical demands, and the implications of these inside and outside of hospitals;

(b) inform its own research personnel of these principles and of the ethical expectation which it and society have of researchers and research institutions;

(c) encourage hospitals to inform patients of the nature of their normal research needs and procedures, and in turn to encourage patients to identify themselves with those needs and procedures. The University might compose an appropriate "statement" for hospitals to use.

(d) exercise strict supervision of research procedures and ethical protocol on the part of those under its auspices, with particular reference to minimizing the threat to privacy; maintaining respect, courtesy and sensitivity towards patients; avoiding unnecessary or trivial research and research procedures; observing the regulations maintained by the hospital; and erring on the side of seeking explicit consent where possible, bearing in mind the difference between research through the use of patient records and experimentation, which involves the "use" of the patients themselves.

Source: University of Toronto Bulletin, April 22, 1977.

ARF has in place a two-tier system of reviewing research proposals that use ARF data or data acquired from outside the Foundation. In order to assure bona fide research, ARF's Research Review System has a series of standing committees which assess from a scientific perspective all research proposals which involve ARF patients, records, staff, or other resources. After the first step assesses the scientific validity of the particular research project, research proposals are reviewed by the Ethics Committee of ARF. This second level of screening of proposed access to medical records for research is designed in part to determine whether the research is "aimed at the public good."

The Foundation has recently sponsored an ad hoc Committee on Ethics, which produced draft "Ethical Guidelines for Research" between September, 1977 and June, 1978. These draft Guidelines are intended to represent the policy, practice, and underlying philosophy of the Foundation. Since the Guidelines are still in draft form, they cannot be quoted in detail here; however, this writer has had an opportunity to review the draft and the various detailed minutes of the ad hoc committee. In their current form the draft Guidelines are eighteen pages in length. Table XVIII is a table of contents, which indicates the range of topics included and discussed. The chairman of both the ad hoc committee and the regular Ethics Committee is Professor Gordon Watson, who is director of the Centre of Criminology at the

TABLE XVIII: TABLE OF CONTENTS
"ETHICAL GUIDELINES FOR RESEARCH"
ALCOHOLISM AND DRUG ADDICTION
RESEARCH FOUNDATION

University of Toronto. Several other members of the ad hoc committee also came from outside ARF. It is also useful to note that the ad hoc committee began its deliberations with a review of the University of Toronto's Handbook on the Use of Human Subjects.[9] As Table XVIII makes clear, the draft Guidelines address a substantial number of ethical issues that are not directly relevant to research and statistical uses of personal data; these will not be discussed here.

The goal of ARF's Ethics Committee is to establish and implement policy concerning the ethical aspects of research involving human subjects and/or non-public sources of information about humans.[10] Its goal is to serve as a source of objective assessment of proposed research projects from an ethical point of view. By having Foundation staff and outsiders on the committee it is intended to obtain some objectivity and distance from the actual research task. The committee is prepared to engage in prior consultation with researchers and to assist them in legitimately overcoming ethical problems. The basic principles followed by the Ethics Committee obviously include a strong

9 V.C. Matsubara, ed., Handbook on the Use of Human Subjects (University of Toronto, Toronto, 1975).

10 This discussion is based on the draft "Ethical Guidelines for Research." (June, 1978).

concern for the personal privacy of individuals and the confidentiality of data pertaining to them. Patients entering an ARF treatment facility are given notice of the research interests of the Foundation and the right to object to participation. Another basic practical principle is that data routinely collected for research purposes should be used for research whenever possible instead of placing a new burden on respondents.

The draft Ethical Guidelines include specific provisions for the handling of research data in order to protect the confidentiality of patients and subjects of research. The records of patients and research subjects are only available for research and statistical purposes with the specific consent of the individual involved. Researchers are expected to sign a declaration to maintain the confidentiality of data, which is "to be broken in only the most extreme of circumstances." This phrase probably refers to the threat of a contempt of court proceeding against a researcher, who has refused to produce data. The Guidelines also point out that the procedures in force to protect confidentiality of research data must be correlated with the sensitivity of the particular data. Researchers should collect identifiable data only when necessary for a particular project; data should be anonymized as soon as possible thereafter.

Extra precautions such as the use of randomized response techniques and the introduction of random errors into the data should be used for extremely sensitive data. Researchers are also expected to inform respondents of the risk of subpoena of research data. Precautions are particularly important when data are collected in identifiable form for longitudinal studies. Conditions for re-contact with patients or research subjects must also be established in this particular case.

It is possible for outside researchers to obtain access to ARF data, but there have been no requests by outsiders reviewed by the regular Ethics Committee during the last year. There are suggestions that most such requests are not pursued by researchers, such as graduate students preparing doctoral dissertations, because of the series of legitimate hurdles placed in the way of outsiders seeking access of this type, including the submission of a formal proposal and review by the Ethics Committee.

The public should also be aware that it is ARF policy to report situations to the police where it is believed that a serious danger to the public is about to occur. This might involve an aggressive patient threatening to murder a spouse. Such a report to the police would only be made after consultation with the Director of the Clinical Institute. ARF does not report to

the police any retrospective information collected on illegal or criminal activity. Thus the general policy applies only to patient-related information from the Clinical Institute; the issue has never arisen in the context of research data.

Uses of ARF Data

The research staff of ARF make substantial use of the personal data collected by the Foundation in its research and treatment activities. The above discussion of "Sources of Data" has already made reference to a number of such studies. The most relevant research categories in the latest Annual Report of the Foundation concerned the areas of epidemiology and evaluative and treatment research. Epidemiological research comprised "studies to discover the prevalence of alcohol and other drug problems in the population at large, and factors causing changes in prevalence." Evaluative and treatment research comprised "studies to assess the effectiveness of prevention or treatment."[11] Specific topics studied under each of these two categories included the following:

11 ARF, Annual Report, 1976-77, p. 9.

Epidemiological Research

Morbidity studies of the alcoholic population.
Mortality studies of the alcoholic population.
Epidemiology of hepatitis B in drug abusers, chronic alcoholics, and ARF staff.
Analytical studies of street drugs.
Clinical epidemiological survey of drug overdose patients.
Studies of the social and behavioral consequences of benzodiazepine use.
Surveys of treatment resources and the prevalence of alcohol and drug problems.
A study of alcohol problems in northwestern Ontario.
Ontario school surveys.
Studies concerned with the forecasting of alcohol consumption in Ontario.

Evaluative Research

Studies of intervention aimed at young people.
Studies of interventions aimed at the drinking driver.
Studies of interventions aimed at the employed alcoholic.
Studies of interventions aimed at the general public.[12]

It is some measure of ARF research activities that in 1976 the ARF staff published 152 publications, including articles, theses, book chapters, and books.[13]

It is perhaps appropriate to mention that the substantial ethical concerns in force at ARF have aroused some criticism from the ARF research staff. There are allegations that the Ethics Committee places a heavy burden on researchers by making access to ARF records so difficult. Although ARF is admittedly a research organization, it is alleged that the Ethics Committee has been timid, conservative, and legalistic in its decision making.

12 Ibid., pp. 14-20.

13 See the bibliography in ibid., pp. 57-68.

Such complaints are perhaps inevitable in a period when procedures for the promotion of ethics, confidentiality, and data protection have been substantially formalized. Other researchers within ARF mention that the concept of informed consent hardly existed at ARF before recent times. There is also substantial sensitivity among some researchers that there is a social stigma attached to treatment by ARF and that, for example, some patients may not want to participate in follow-up studies. Continued representation of researchers on the Ethics Committee should help to alleviate some of these criticisms over time.

Recommendations

ARF is to be congratulated for having produced excellent guidelines for the protection of patient data in research and treatment activities. In particular the ad hoc Ethics Committee has provided researchers with relatively clear and flexible guidelines for the future. At least with respect to protecting the confidentiality of personal data, these guidelines leave nothing to be desired. They should be approved and implemented by the Foundation as soon as possible. ARF should continue to use a blanket consent form whereby patients and other research subjects can agree in advance that data concerning them and

already furnished by them can be used for research and statistical uses in the future; a research and statistical use of data implies that these will not be used to directly affect an individual, unless treatment is built into the research protocol, as in the therapeutic research conducted by ARF. In the same connection the Research Division at ARF should maintain a strict functional separation between administrative or treatment uses of data and research and statistical uses of data. This concept has already been discussed above in Chapter III in connection with the Ministry of Health. Research protocols should continue to include written declarations or undertakings concerning the uses of data in a particular research project. The draft Guidelines are perhaps a bit weak in lacking specific measures whereby research projects can be monitored for compliance with their written undertakings; while this task may not be appropriate for the Ethics Committee, the administration of ARF should institute appropriate monitoring procedures as a regular part of administration.

In The Patient's Handbook. A Guide to the Clinical Institute of the Addiction Research Foundation (1978), there are three paragraphs on the use of patients and their records for research. Each patient receives a copy. This admirable practice should also be implemented for other research subjects of the Foundation, describing in simple detail ARF's concern for

ethics, privacy, and confidentiality, and the practices that are in force for these purposes. A simple leaflet may be appropriate for this purpose. As a wise British statistician said recently, there is "no use doing good by stealth." Patients and research subjects at ARF should both receive reasonable explanations about how information supplied by them to the Foundation will or may be used for research purposes.

Because of the sensitive subjects studied by the Foundation, it has a particular responsibility to develop "ethically-alert" researchers. This can occur just as the Foundation has to avoid imposing too much of a bureaucratic burden on its own researchers; research is difficult enough to accomplish without adding unnecessary burdens. A recent statement by the American privacy expert, Alan F. Westin, is worth repeating: "There is no reported U.S. case in which a physician or hospital had to compensate a patient for an injury resulting from breach of confidentiality."[14] Similarly, there are few documented abuses of privacy or confidentiality by American researchers. Another privacy expert, Arthur Miller of the Harvard Law School, and Mary Kane of the State University of New York Law School at Buffalo, found in a review of American research organizations in

14 U.S. Privacy Protection Study Commission, Personal Privacy in an Information Society. Report of the Privacy Protection Study Commission (Washington, D.C., U.S. Government Printing Office, July, 1977), p. 284.

the early 1970s that researchers were primarily guilty of insensitivity to privacy and confidentiality.[15] It is most evident from daily experience that researchers in the health and social sciences resent regulation. Yet some regulation in the public interest is essential; this must be supplemented by self-regulation in the researchers' own interest. The burden is especially heavy on researchers using sensitive data at an institution like ARF. Fortunately, the same researchers are in an excellent position to recognize when the sensitivity of data collected or used by them requires extraordinary statistical or technical protective measures. ARF might particularly consider some of the new automatic devices for scrambling or encrypting personal data, which are available from commercial manufacturers at a reasonable cost.[16]

Relatively simple legal changes might also be considered in order to promote public confidence in the integrity of ARF and its staff. This would primarily involve amendments to the Alcoholism and Drug Addiction Foundation Act of 1965. Although ARF is currently subject to the Public Hospitals Act and

15 See David H. Flaherty, Privacy and Government Data Banks. An International Perspective (New York, forthcoming, 1978), chapter 21.

16 See David H. Flaherty, "Privacy, Confidentiality, and Security in a Canadian Electronic Funds Transfer System," (Ontario Ministry of the Attorney-General, Toronto, 1978), pp. 32-40, 61-67.

accompanying regulations, this occurred in the early 1970s in order to secure federal funding under the OHIP system. Since the format of federal funding has ceased, it would probably be useful for the Foundation to operate fully under its own statute. This could mean that certain types of extremely sensitive data might no longer be furnished to MOH or HMRI. The current ARF Act does not mention or discuss secrecy, confidentiality, or sanctions for breach of either of these values. A specific provision on confidentiality comparable to the 1972 amendment to the Cancer Act of the Ontario Cancer Treatment and Research Foundation seems appropriate. It would thus be possible to turn the Foundation into a data enclave, whereby identifiable data would never leave the Foundation for any purpose. This would be comparable to the general situation of national statistical agencies or such institutions as the United States National Center for Health Statistics. Provision could probably be made whereby outside researchers requiring legitimate access to ARF identifiable data could be sworn in under the Foundation's Act as staff members and thus become subject to the provision on confidentiality. It would also be useful from a public relations' point of view to subject ARF staff to criminal sanctions for breach of secrecy and confidentiality.

Amendments to the ARF governing statute should also address the issue of the desirability of granting ARF researchers immunity from legal subpoena for their research data. Researchers could be granted by legislation a privilege under which they would be rendered immune from subpoena or other formal legal requests for information generated in the course of research projects. Such protection would be especially useful for longitudinal projects that collect continuing data on such sensitive topics as the long-term impact on behavior of drinking practices. It is asserted that the general situation at ARF at present, whereby data could in theory be subpoened, has a chilling effect on potential research projects. It is also alleged that the Ethics Committee has discouraged legitimate research projects that would involve data collection on illegal or criminal activities. It may be more difficult to fashion such a legal privilege for the Clinical Institute Division of ARF.

On the whole, this question of legal privilege for research data is a very thorny issue that remains unresolved in most jurisdictions.[17] However, an Ontario model may now be available. On December 13, 1977, the Minister of Health introduced in the Ontario legislature proposed amendments to the Mental Health Act,

17 See Paul Nejelski, ed., Social Research in Conflict with Law and Ethics (Ballinger Publishing Company, Cambridge, Massachusetts, 1976).

which would allow for the recognition of the concept of "medical privilege" in court cases in which hospital psychiatric records on individuals are subpoened. Ontario could become the first Canadian province to recognize this particular medical privilege.[18] Under these particular amendments, lawyers would have the right to argue before a court that particular psychiatric records should not be disclosed. As Dennis Timbrell, the Minister of Health, said in the legislature in explaining the above provisions: "The great value of medical privilege lies in the inviolable nature of medical confidences, recognized by law and secure from controversy and interference. Legislative action will now bring recognition of such privileges."[19]

18 Globe and Mail, December 14, 1977, p. 5.

19 Ibid., p. 5.

GENERAL RECOMMENDATIONS

Since specific recommendations have been presented above in connection with each of the various agencies under review, only a few general issues require consideration at this point, including the right of researchers to obtain access to government data for scientific purposes, and the necessity of a regulatory apparatus to control research and statistical uses of personal data in the hands of the government.

1. The Ontario government should explicitly recognize the legitimacy of using the personal data of its ministries and agencies for approved research and statistical purposes.

Researchers should generally be entitled to use government data for scientific purposes under controlled conditions. This general recommendation has specific relevance to the important field of health research that has been reviewed in this Working Paper. In this connection Professor Alan F. Westin of Columbia University, the leading American privacy expert, recently devoted a specific section of his report on computers and medical records to "the importance of research and evaluation using health data."[1]

1 Alan F. Westin, Computers, Health Records and Citizens' Rights (National Bureau of Standards, U.S. Department of Commerce, Washington, D.C., December, 1976), pp. 300-303.

Government agencies should set out the right of approved researchers to obtain access to their data, and the general criteria for access, in general policy statements, statutes, and regulations. Any current legal provisions that hinder access to personal data for approved health research should also be changed to recognize the legitimacy of such scientific uses of government data. Government agencies should be encouraged to make their policies on data dissemination generally known to the research community. Proposals for new legislation and regulations affecting particular agencies of government should also be reviewed with respect to their potential impact on research and statistical uses of personal data.

The West German state of Hesse was the first jurisdiction in the Western world to enact data protection legislation in 1970. The 1978 revision to the Hesse Data Protection Act provides a valuable model for recognizing the legitimacy of research uses of government data. This initiative was partly in response to the enactment of the first Federal Data Protection Law in the Federal Republic of Germany in January, 1977. Professor Spiros Simitis, the current Hesse Data Protection Commissioner, was one of those who pointed out that the new federal law did not furnish a clear legal basis for research with government personal data. Simitis argues that although persons require data protection, research must still be possible. In the January 31, 1978

revision of the Hesse Data Protection Law, Simitis and others were successful in obtaining a provision to authorize state agencies to disseminate data for research purposes. Section 15 of the new law authorizes "data processing for scientific purposes:"

> "Universities and other public institutions with the task of carrying out independent scientific research, can store and alter personal data for specific research projects in the scope of their competence: for this purpose the authorities and public institutions mentioned in section 3, subsection 1, may release personal data to them. Data processing as described in the foregoing sentence is permissible only on condition that the persons hereby concerned have given their consent, or if their vested interests are not harmed or involved because of 1) the type of data, 2) general public knowledge of the data, or 3) the type of data processing employed.
>
> Further communication of the personal data stored, altered or communicated according to subsection 1 is permitted only with the authorization of the persons hereby concerned."[2]

This admirable provision was in part the product of formal and informal pressure from social scientists and other researchers, who are now seeking a similar provision in other forthcoming data protection laws for the Lander and in a future revision of the federal law. The new Hesse provision does not compel the disclosure of data for research and statistical purposes, but it does remove any legal barriers to such releases. The provision should be especially useful in obtaining access to

2 This unofficial translation is from David H. Flaherty, Privacy and Government Data Banks. An International Perspective (Science Associates/International, Inc., New York, 1978, forthcoming), chapter 9.

state administrative data. Data for release under section 15 can include both identifiable and anonymized microdata.

The explicit recognition of the legitimacy of using Ontario government data for research does not in any way imply that the actual process of obtaining access should be unregulated. The government agencies in question should continue to have advisory committees for scientific and ethical review of research proposals. Questions like cost recovery and the burden on current staff must be given adequate consideration.[3]

Although it is not within the terms of reference of this Working Paper to engage in detailed prescriptions for the research community, a few words are in order. An unnecessarily large number of researchers have yet to learn that obtaining access to any personal data in the hands of government for scientific purposes requires sensitivity, tolerance, and a capacity for accommodation. Strident approaches to government custodians will continue to have a low level of success, even when the legitimacy of research uses of government data is recognized.

3 For an informed discussion of the conditions under which a government custodian should release data, especially for biomedical and epidemiological research, see Personal Privacy in an Information Society. The Report of the Privacy Protection Study Commission, Washington, D.C., 1977, pp. 306-310.

Researchers have a further responsibility to keep themselves and their associates in an ethically-alert posture with respect to the use of data. Researchers have to share accountability with the custodians of personal data for the protection of confidentiality and the interests of research subjects. Every researcher in Canada should be aware of two major recent reports on ethics.[4] Finally, researchers using Ontario government data should make use of their own professional associations to articulate their research needs to provincial agencies. This will particularly assist the latter in formulating methods of disseminating data that might be of considerable use to as many researchers as possible.

2. The Ontario government should create the office of Data Protection Commissioner with, among other duties, the responsibility to promote, regulate, and monitor scientific uses of government data by the research community.

The current practices and procedures of the government agencies in the health field reviewed in this Working Paper do not indicate a need for a strong regulatory or licensing mechanism,

4 Canada Council, Ethics. Report of the Consultative Group on Ethics (Ottawa, 1977); and Medical Research Council, Ethics in Human Experimentation, Report No. 6 (Ottawa, 1978).

such as the Data Inspection Board in Sweden. Under the Swedish Data Act of 1973, the Data Inspection Board has to license any computerized personal data files in the public or private sectors.

The model of the Hesse Data Protection Act of 1970 would seem more appropriate to the province of Ontario today. The Hesse Data Protection Commissioner has the right to screen and investigate particular government data banks and to promulgate codes of fair information practices. But this official does not have the direct power to order the correction of abuses. His recourse is to the government in power, to the legislature, and to public opinion in order to rectify an alleged abuse. As a creature of the legislature, he prepares an annual report to it. Because of their recourse to public opinion and to the legislature, the successive Hesse Data Protection Commissioners (and their small staffs) have been very successful during the 1970s in achieving data protection in the state of Hesse, which has a population of approximately 6 million persons.

One of the advantages of the Hesse model for Ontario is that substantial segments of the research community could be indirectly regulated by monitoring the activities of government custodians of data. This coverage should include such publicly-funded research organizations as the Ontario Economic Council, the Ontario Cancer Treatment and Research Foundation, and the Alcoholism and Drug Addiction Research Foundation.

An Ontario Data Protection Commissioner could have a variety of important tasks to perform in connection with research and statistical uses of government data. In particular, this official could make certain by investigations that government custodians abide by the rules that they have promulgated for data protection and data dissemination. The responsibilities of such a Commissioner could also include issues such as the following: ensuring that data collected for statistical purposes are only used for such purposes; receiving complaints from the general public or from government custodians that researchers may have abused their privileges of access; establishing the liability of custodians and researchers for lack of reasonable care to protect confidentiality; reviewing the necessity for government custodians to operate their own dedicated computers in the handling of sensitive information; and evaluating the necessity of such protective practices as the encoding and/or encrypting of stored data.

The duties of the Data Protection Commissioner should be defined broadly enough to include responsibilites to facilitate access to data for legitimate scientific purposes. This might necessitate a change of title for this individual to a description such as the Fair Information Practices Commissioner. Researchers requiring access to identifiable government data for epidemiological research could seek the approval in advance of this official. A government agency could require a researcher

to have a proposal reviewed by this person as well. Government agencies could approach the Data Protection Commissioner with problems concerning access to data and confidentiality, especially with respect to such sensitive issues as the uses of Unique Personal Identifiers, record linkages, and the creation of data enclaves. In such areas the Commissioner could make a particularly effective contribution by monitoring existing practices on a regular basis and promoting the imposition of sanctions for data abuses. If a particular research project had received government data under certain conditions, the Commissioner and his associates would have the right to visit the particular project in question to ensure that the undertakings were being complied with. An official with these types of responsibilites could promote the interests of both the general public and the research community.

APPENDIX I

Final Report of the Bellagio Conference on Privacy, Confidentiality, and the Use of Government Microdata for Research and Statistical Purposes

DAVID H. FLAHERTY
Conference Chairman and Organizer

Introduction

For the last three years the Privacy Project of the University of Western Ontario has been studying the problems of privacy and confidentiality involved in the collection and dissemination of individual microdata by central statistical agencies in Canada, the United States, the Federal Republic of Germany, Sweden, and the United Kingdom. The focus has included descriptions of the personal data collected, as well as the process of formulating policy for data protection and dissemination. The Privacy Project was originally funded in the Ford Foundation's Competition on the Common Problems of Advanced Industrial Societies with David H. Flaherty and Edward H. Hanis as co-principal investigators.

The custodians of individual data (microdata) collected by governments for statistical purposes do maintain a very high level of concern for the protection of confidentiality. In fact, a basic problem for researchers is to persuade officials in central statistical agencies that they should be doing more, within the boundaries of maintaining confidentiality, to satisfy demands for access to data for research and statistical purposes. The publication of aggregated statistics and the provision of special tabulations is not satisfying the needs of researchers, in and out of government, in a wide variety of fields, including demography, sociology, economics, medical research, and planning of various sorts. The provision of public use samples from censuses of population and social surveys, which is currently undertaken in the United States and Canada, is only one of the ways in which researchers' needs for access to microdata, in many cases in anonymized form, can be more adequately satisfied.

Through a process of intensive interviewing in each country, the Privacy Project identified key personnel in each central statistical agency, who are directly involved in the formulation of policy on confidentiality and data dissemination. They face common problems in fashioning ways to protect the confidentiality of data they release to the general public as well as researchers. Yet the principal decision-makers from these several countries have had surprisingly little contact with one another on these specific issues. The Bellagio Conference brought together leading figures from each of the five statistical agencies for a thorough discussion of common problems and solutions.

Since it is necessary to demonstrate to custodians of government data that a real need exists for access to microdata for research and statistical purposes, the Bellagio Conference included participants who are leading exponents of the use of microdata from each country. These researchers were invited to discuss the uses of microdata based on their own activities and other major research undertakings with which they are familiar.

Five national case studies prepared by David H. Flaherty, served as background reading for discussions at the Conference of the current situation in each country. These case studies will be published in a volume tentatively entitled "Privacy and Government Data Banks. An In-

Note.—This Conference was held at the Rockefeller Foundation's Bellagio Study and Conference Center, Bellagio, Lake Como, Italy, August 16–20, 1977. The report is reprinted here with the permission of Mr. Flaherty.

Source: Statistical Reporter, No. 78-8 (May, 1978), 274-79. (Office of Federal Statistical Policy and Standards, Washington, D.C.)

ternational Perspective on Research and Statistical Uses of Government Data" (Science Associates/International, Inc., New York, forthcoming, 1978).

The invited participants in the Conference, who are listed below, met for a total of eighteen hours of formal sessions from August 16 to 20, 1977. Since the participants lived together at the Villa Serbelloni in Bellagio for the period of the Conference, there were many more hours of informal discussions. After a general introductory session on August 16 in which participants described their particular interests in the topic of the Conference, there were separate sessions on August 17 and 18 devoted to a consideration of the specific situation in the United Kingdom, Canada, Sweden, the United States, and the Federal Republic of Germany. Although these discussions of specific countries were wide-ranging in character, a number of basic themes began to emerge.

On August 19 and 20 there were a number of general sessions which had evolved on the basis of previous discussions and specific suggestions from participants. At the outset of these general discussions the group decided to formulate a series of general principles or propositions pertaining to research and statistical uses of data held by government agencies. A series of eighteen general Bellagio Principles subsequently evolved on the basis of group discussions. The Principles are presented below in the sequence in which they emerged and were debated. They should be interpreted collectively, since general statements are frequently qualified in subsequent principles. None of the Principles should be attributed to any one participant, nor are they binding on individual participants in the Conference. Nevertheless, it was agreed that the Principles would be made available in this *Report* of the Bellagio Conference, which has been reviewed by all participants.

The Bellagio Principles: A Discussion

1. *National statistical offices should provide researchers both inside and outside government with the broadest practicable access to information within the bounds of accepted notions of privacy and legal requirements to preserve confidentiality.*

Discussion:

The participants substantially agreed on the importance of facilitating access to government data for research and statistical purposes. They thought it was important to establish this basic trust toward openness, since researchers perform a public service. It was understood that national agencies have to allocate scarce resources between researchers and other public users.

2. *Legal and social constraints on the dissemination of microdata are appropriate when they reflect the interests of respondents and the general public in an equitable manner. These constraints should be re-examined when they result in the protection of vested interests, or the failure to disseminate information for statistical and research purposes (i.e., without direct consequences for a specific individual).*

Discussion:

The participants recognized that data dissemination could not take place in an unrestricted manner because of concern for privacy and legal considerations. At the same time the public has a substantial interest in beneficial uses of data. The participants thought that some of the current reasons for restricting access to data did not adequately reflect the public interest in data use.

3. *All copies of government data collected or used for statistical purposes should be rendered immune from compulsory legal process by statute.*

Discussion:

This particular statement reflected the consensus that individual data used for research and statistical purposes should not directly result in decisions about a particular individual, unless the research project specifically involved such a result, as in some research in the drug treatment field.

4. *In making data available to researchers national statistical offices should provide some means to ensure that decisions on selective access are subject to independent review and appeals.*

Discussion:

The participants agreed that access to government information for research and statistical purposes should be non-discriminatory. In particular, decisions on access to data should not rest solely on the hands of members of the bureaucracy. If a researcher believes that he or she has a legitimate right of access to data, a means of appealing a negative decision should exist.

5. *The distinction between a research file, in the sense of a statistical record (as defined in the 1977 report of the U.S. Privacy Protection Study Commission), and other micro files is fundamental in discussions of privacy and dissemination of microdata. All dissemination of government microdata discussed in connection with the Bellagio Principles is assumed to be a transfer of data to research files for use exclusively for research and statistical purposes.*

Discussion:

The discussion at Bellagio emphasized the importance of the distinction between research/statistical and administrative/regulatory uses of individual information. The report of the American Privacy Protection Study Commission on *Personal Privacy in an Information Society* defined a research and statistical record as "any item, collection or grouping of information maintained in any form of record *solely* for a research or statistical purpose." Research and statistical data is "not intended to be used, in whole or in part, for making a decision about an individual that is not an integral part of the particular research project." The discussions at Bellagio emphasized the importance of informing data disseminators, respondents, and data users of the distinction.

6. *There are valid and socially-significant fields of research for which access to microdata is indispensible. Statistical agencies are one of the prime sources of government microdata.*

Discussion:

After considerable debate and a process of mutual education in the early stages of the Bellagio Conference, the participants concluded that a number of important fields of research required access to microdata, especially from government sources. It was also recognized that not every researcher requires access to microdata either in anonymized or identifiable form. There was a further sense that statistical agencies in particular should collect statistics on the utilization of microdata in order to justify the efforts and risks involved.

7. *Public use samples of anonymized individual data are one of the most useful ways of disseminating microdata for research and statistical purposes.*

8. *Techniques now exist that permit preparation of public use samples of value for research purposes within the constraints imposed by the need for confidentiality. Countries with strict statutes on confidentiality have prepared public use samples.*

Discussion:

There was a general consensus that public use samples were the single most important way in which central statistical agencies could assist the research community. The discussion emphasized the more than fifteen years experience with public use samples in the United States without any untoward results in terms of detriment to an individual. Public use samples provide researchers with access to microdata at low cost and with great flexibility for their research needs.

9. *There are legitimate research purposes requiring the use of individual data for which public use samples are inadequate.*

10. *There are legitimate research uses which require the utilization of identifiable data within the framework of concern for confidentiality.*

Discussion:

The participants wished to emphasize the legitimacy of research and statistical activities involving the need for access to identifiable individual information. They also recognized that such sensitive uses of data require substantial justification in terms of potential payoffs to the general public. Nevertheless, data protection should be implemented in such a fashion as to make possible the conduct of such important types of research as longitudinal studies.

11. *Other techniques of extending to approved research the same rights and obligations of access enjoyed by officers of the government agency need to be considered in terms of better access.*

Discussion:

The participants concluded that researchers share with government statisticians a high level of concern for the confidentiality of personal information. At the same time, non-governmental researchers do not normally enjoy the same level of access to individual data enjoyed by employees of a government agency. Within the framework of Principle No. 1, the Conference concluded that various means should be explored to improve access for research and statistical purposes, so as to improve on existing uses of government information. In particular, the participants focused on the possibility of swearing-in researchers for a tempor-

ary period as employees of a statistical agency. Existing examples of mutual cooperation between a statistical agency and researchers were applauded.

12. There is considerable potential for development of more economical and responsive customized-user services, such as 1) record linkage under the protection of the statistical office, 2) special tabulations, 3) public use samples for special purposes. Such services must often involve some form of cost recovery.

Discussion:

Even if one takes into account the broad mandates and heavy current burdens of statistical agencies, more can be done to assist research and statistical users of data from outside the agency. The discussion emphasized the technical capacity of statistical agencies to implement in a more economical and responsive manner the three forms of services mentioned in Principle 12. Contemporary restraints on the budgets of government agencies emphasize the importance of cost recovery for the agency in connection with such forms of dissemination.

13. Some research and statistical activities require the linking of individual data for research and statistical purposes. The methods that have been developed to permit record linkage without violating law or social custom regarding privacy should be used whenever possible.

Discussion:

The participants recognized that the linkage of individual data, even for research and statistical purposes, was one of the most sensitive uses of personal information, particularly in the minds of the general public. Nevertheless, the Conference concluded that it was important to make a general statement recognizing the legitimacy of certain forms of record linkages, such as in various types of medical and epidemiological research. It would not be in the public interest absolutely to prohibit the linkage of personal records for research and statistical purposes.

14. Professional or national organizations should have codes of ethics for their disciplines concerning the utilization of individual data for research and statistical purposes. Such ethical codes should furnish mutually agreeable standards of behavior governing relations between providers and users of governmental data.

15. Users of microdata should be required to sign written undertakings for the protection of confidentiality.

Discussion:

After considerable emphasis on the responsibility of government agencies to improve data dissemination, the Conference turned to the counter-issue of encouraging researchers to share accountability for data protection with government agencies. Despite long traditions of concern for confidentiality within particular disciplines in the social sciences, individual researchers must be continually sensitized to the need for care and caution with respect to the protection of confidential information. There was overwhelming agreement on the desirability of up-to-date ethical codes and the utilization of written undertakings to govern data dissemination.

16. Considerable efforts should be made to explain to the general public the procedures in force for the protection of the confidentiality of microdata collected and disseminated for research and statistical purposes.

Discussion:

The Conference felt that there is considerable public misunderstanding of the special nature of data use involved in research and statistical activities. Given widespread suspicions of the abuse of administrative data by government agencies, it is difficult to convince the general public of the long-standing concern for confidentiality of statistical agencies in particular. Yet, however demanding, such efforts at public relations on the part of statistical agencies must be continued and improved. Certain statistical agencies have made considerable advances in improving general contacts with the public, including attempts to explain the purposes for which data is being collected. The Conference strongly supported the notion of informed consent in data collection and use, whenever practicable. Since individual consent is not always desirable or practical, such as in certain types of medical research, some form of surrogate has to be considered under special circumstances.

17. The right of privacy is evolving rather than static, and closely related to how statistics and research are perceived. Therefore, statisticians and researchers have a responsibility to contribute to policy and legal definitions of privacy.

Discussion:

The thrust of this particular Principle is to emphasize that policies on data protection have often been formulated in the past without adequate attention to the specialized nature of research and statistical uses of government data. Both statisticians and researchers have been too hesitant to involve themselves in public debates on these issues and to articulate their particular traditions and concerns. This Principle emphasizes the importance of good public relations on the part of both statisticians and researchers.

18. Public concern about privacy and confidentiality in the collection and utilization of individual data can be addressed in part as follows:

(1) voluntary data collection, whenever practicable;

(2) advanced general notice to respondents and informed consent, whenever practicable;

(3) provisions for public knowledge of data uses;

(4) public education on the distinction between administrative and research uses of information.

Discussion:

The particular concerns of the Bellagio Conference with research and statistical uses of data meant that ways of protecting personal privacy as such were not a central focus of discussion. Nevertheless, the Conference closed with the presentation of a general resolution emphasizing ways in which statistical agencies in particular could attempt to alleviate public concern with problems of privacy and confidentiality. In many ways Principle 18 reiterates general tendencies included in earlier Principles.

CONFERENCE PARTICIPANTS

A. M. Adelstein
Chief Medical Statistician
Office of Population Censuses and Surveys
United Kingdom

Lois Alexander
Privacy Coordinator
Office of Research and Statistics
Social Security Administration
Department of Health, Education, and Welfare
United States

Robert Boruch
President, Council for Applied Social Research
Department of Psychology
Northwestern University
United States

John Bossons
Institute for Policy Analysis
University of Toronto
Canada

Theodore Clemence
Chief, Office of Program and Policy Development
Bureau of the Census
United States

David H. Flaherty
Co-principal Investigator, Privacy Project
Professor of History
The University of Western Ontario
Canada

Michael Francino
Director General
Policy, Planning, and Evaluation
Statistics Canada
Canada

Edward Hanis
Co-principal Investigator, Privacy Project
Director
Social Science Computing Laboratory
The University of Western Ontario
Canada

*Peter E. Hart
Department of Economics
University of Reading
United Kingdom

* unable to attend in person

Carl-Gunnar Janson
Department of Sociology
Stockholm University
Sweden

Peter Kirkham
Chief Statistician
Statistics Canada
Canada

Hans-Jürgen Krupp
Seminar für Sozialpolitik
Johann Wolfgang Goethe-Universität
Federal Republic of Germany

*Edmund Rapaport
Head, Department of Systems and Information
Statistiska Centralbyran
Sweden

Brigitte Reimann
Chief of Section
General and Coordinating Tasks in Official Statistics
Statistisches Bundesamt
Federal Republic of Germany

Grant Reuber
Chairman
Ontario Economic Council
Canada

Erwin Scheuch
Director
Institute for Applied Social Research
University of Colone
Federal Republic of Germany

Eric J. Thompson
Chief Statistician
Social Monitoring Branch
Central Statistical Office
United Kingdom

Klas Wallberg
Director
Department of Statistics viduals
Statistiska Centralbyran
Sweden

Conference Secretary

Peggy Reuber
Weldon Library
The Univesity of Western Ontario
Canada

LIST OF COMMISSION RESEARCH PUBLICATIONS TO DATE

The following list of Research Publications prepared for the Commission may be obtained at the Ontario Government Bookstore in Toronto, or by mail order through the Publications Centre, 880 Bay Street, 5th Floor, Toronto, Ontario M5S 1Z8. Further titles will be announced upon their completion.

1 The Freedom of Information Issue: A Political Analysis
by Donald V. Smiley, Professor of Political Science
York University, Toronto, Ontario

2 Freedom of Information and Ministerial Responsibility
by Kenneth Kernaghan, Professor of Politics and Administration
Brock University, St. Catharines, Ontario

3 Public Access to Government Documents: A Comparative Perspective
by Donald C. Rowat, Professor of Political Science
Carleton University, Ottawa, Ontario

4 Information Access and The Workmen's Compensation Board
by Terence G. Ison, Professor of Law
Queen's University, Kingston, Ontario

5 Research and Statistical Uses of Ontario Government Personal Data
by David H. Flaherty, Professor of History
University of Western Ontario, London, Ontario

ISBN 0-7743-3229-8